

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 21:41:11 ; Search time 9384 Seconds

(without alignments) 11230.094 Million cell updates/sec

Title: US-09-942-429A-6

Perfect score: 2576

Sequence: 1 gaattccgcagagatccagc.....aaaaaaaaaacggaattc 2576

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank1: *
1: gb_ba: *
2: gb_hhg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sv: *
13: gb_un: *
14: gb_vt: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vt: *
30: em_hhg_hum: *
31: em_hhg_inv: *
32: em_hhg_other: *
33: em_hhg_mus: *
34: em_hhg_pln: *
35: em_hhg_rnd: *
36: em_hhg_mam: *
37: em_hhg_vit: *
38: em_ay: *
39: em_hhgo_hum: *
40: em_hhgo_mus: *
41: em_hhgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2576	100.0	2576	6	AX336419
2	2576	100.0	2576	6	AX409491
3	2576	100.0	2576	6	HMCYTFPAOH
4	2417.2	93.8	2470	9	S67580
5	2394	92.9	2469	9	S67581
6	2342.6	90.9	2382	9	HUMRPAOH
7	1625.8	63.1	1763	6	AX336099
8	1625.8	63.1	1763	6	AX336427
9	1625.8	63.1	1763	6	AX336709
10	1625.8	63.1	1763	6	AX408963
11	1625.8	63.1	1763	9	HUMOMHY
12	1389.4	53.9	1833	9	BC041158
13	1241.4	48.2	1790	4	RABCTPA46
14	1219	47.3	2191	4	RABCTYR9
15	1197.2	46.5	2377	4	OCF450XD
16	1181.2	45.9	2356	4	RABCTPA45
17	1179.2	45.8	2929	4	RABCTP2
18	1137.8	44.2	2400	4	SGC278474
19	1132.8	44.0	2376	4	RABCTYR4
20	1127.4	43.8	1694	4	RABCTPA47
21	1122.2	43.6	2455	4	AF384031
22	1101.8	42.8	123921	9	AF731892
23	1092.6	42.4	14995	9	AF208532
24	1091.2	42.4	2376	10	AK098123
25	1091	42.4	122432	2	ALJ390073
26	1091	42.4	193358	9	AL135960
27	1091	42.4	193471	9	HS131016
28	1083.6	42.1	1945	18	BC025936
29	1083.6	42.1	2382	10	BC014721
30	1083.6	42.1	2431	10	BC026582
31	1076.8	41.8	1931	10	BC033924
32	1072.6	41.6	2357	10	BC031141
33	1059.2	41.1	1957	10	RATCYPA45OB
34	1050.6	40.8	1515	4	SSC318096
35	1050.6	40.8	1515	4	SSC318097
36	1022.4	39.7	2134	10	RATCYT452
37	1022.4	39.7	2462	6	AX401946
38	1022.4	39.7	2462	10	RATCYPLA
39	1012.8	39.3	2081	10	AK098101
40	1012.2	39.3	1887	10	MMCTP4A
41	1012	39.3	2077	10	AB018421
42	1012	39.3	2083	10	BC010747
43	979.6	38.0	2140	10	BC051049
44	968.8	37.6	2491	10	RATIVAA
45	968.8	37.6	2491	10	AK098088

ALIGNMENTS

RESULT 1
AX336419
LOCUS AX336419 2576 bp DNA
DEFINITION Sequence 6928 from Patent WO0194629.
ACCESSION AX336419
VERSION AX336419.1 GI:18127138
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
1. Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.,
Cancer gene determination and therapeutic screening using signature

Pred. No. is the number of results predicted by chance to have a

Gene sets
Patent: WO 0194629-A 6928 13-DEC-2001;
Avalon Pharmaceuticals (US)
Location/Qualifiers
1. 2576
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 572 a 793 c 583 g 628 t
ORIGIN

Query Match 100.0%; Score 2576; DB 6; Length 2576;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGCGAGATCCAGCAGGTGTGACCAATGATGATCTGTGCTGAGGCCAGCA 60
1 GAATTCGCGAGATCCAGCAGGTGTGACCAATGATGATCTGTGCTGAGGCCAGCA 60
61 GACTCTGGGTGATGCTCTGGAATCTCCAGCGGCTCCCTGCTCATTTGCTTCTGC 120
61 GACTCTGGGTGATGCTCTGGAATCTCCAGCGGCTCCCTGCTCATTTGCTTCTGC 120
121 TGTCTATCAAGGCAAGTTCAGCTCTACCTGCAAGGCAAGTGTGCTCTCAAAAGCTTCAGC 180
121 TGTCTATCAAGGCAAGTTCAGCTCTACCTGCAAGGCAAGTGTGCTCTCAAAAGCTTCAGC 180
181 AGTTCCTGGGCTCCCTCCGCACTGGCTCTTGGGGCAATCCAGAGGCTCCAAAGAGCC 240
181 AGTTCCTGGGCTCCCTCCGCACTGGCTCTTGGGGCAATCCAGAGGCTCCAAAGAGCC 240
241 AGGAGCTACACGAGATTCAGAAATGGGTGAGACATTCCTCAAGTGCCTGCTCATTTGAGC 300
241 AGGAGCTACACGAGATTCAGAAATGGGTGAGACATTCCTCAAGTGCCTGCTCATTTGAGC 300
301 TATGGGAGGCAAAAGTTCGTGTCAGCTCTATGACCTGATATGAAAGTATTTCTGG 360
301 TATGGGAGGCAAAAGTTCGTGTCAGCTCTATGACCTGATATGAAAGTATTTCTGG 360
361 GGAGATCAGACCCGAAATCCATGATGTTCTTACAGATTCCTGGCTCCATGAGATTGGGTAGC 420
361 GGAGATCAGACCCGAAATCCATGATGTTCTTACAGATTCCTGGCTCCATGAGATTGGGTAGC 420
421 GCTTGCTCTGTGAAATGGGACAGATGTTCCAGATGCAAGATGCTGACCCAGCTT 480
421 GCTTGCTCTGTGAAATGGGACAGATGTTCCAGATGCAAGATGCTGACCCAGCTT 480
481 TCCACTATGACATCTGAAAGCCCTATGTGGGCTCATGGCACTCTGTACAGATATGC 540
481 TCCACTATGACATCTGAAAGCCCTATGTGGGCTCATGGCACTCTGTACAGATATGC 540
541 TGGACAAATGGGAAGGCTCTTGGCCAGATTCCTCTGAGAGTCTTTACAGAGTCT 600
541 TGGACAAATGGGAAGGCTCTTGGCCAGATTCCTCTGAGAGTCTTTACAGAGTCT 600
601 CCTTGAAGACCTGGAACCATATGATGATGATGCTTCAAGCATCAAGGAGCATCAAG 660
601 CCTTGAAGACCTGGAACCATATGATGATGATGCTTCAAGCATCAAGGAGCATCAAG 660
661 TGGACAGAAATCTCAGTCTCTATACAGGCTATGATGATGATGATGATGATGATGAT 720
661 TGGACAGAAATCTCAGTCTCTATACAGGCTATGATGATGATGATGATGATGATGAT 720
721 CCCGTGTGAGAAATGCTTTACCAAGATGACATCTACCTGACCTTCTGCTGAGCC 780
721 CCCGTGTGAGAAATGCTTTACCAAGATGACATCTACCTGACCTTCTGCTGAGCC 780
781 GCTGGAACAACGCGGCTGCAAGATGAGGAGCTGAGAAATCAAGAGAGGCAATTTGG 840
781 GCTGGAACAACGCGGCTGCAAGATGAGGAGCTGAGAAATCAAGAGAGGCAATTTGG 840
841 GGAAGCTCAATCAAGAGAGGAGGAGCTGAGAAATCAAGAGAGGCAATTTGG 900
841 GGAAGCTCAATCAAGAGAGGAGGAGCTGAGAAATCAAGAGAGGCAATTTGG 900

841 GGAAGCTCAATCAAGAGAGGAGGAGCTGAGAAATCAAGAGAGGCAATTTGG 900
901 ATTTTCTGATATCTCTCTTGGCCAAATGAGATGAGAGCATTTGTGACAGAG 960
901 ATTTTCTGATATCTCTCTTGGCCAAATGAGATGAGAGCATTTGTGACAGAG 960
961 ACCCTCCGTCTGAGGTGACAGCTTCAATGTTGAGGGCCACGACACACAGCCAGTGGGA 1020
961 ACCCTCCGTCTGAGGTGACAGCTTCAATGTTGAGGGCCACGACACACAGCCAGTGGGA 1020
1021 TCTCCGATCTCTATGCTCTGAGCAACACCCCAAGCATAGAGAGGTGCCGGAGG 1080
1021 TCTCCGATCTCTATGCTCTGAGCAACACCCCAAGCATAGAGAGGTGCCGGAGG 1080
1081 AGATCCACAGCTCTCGGTGATGAGGCTCCATCATCTGAAACCATGAGCAAGATGC 1140
1081 AGATCCACAGCTCTCGGTGATGAGGCTCCATCATCTGAAACCATGAGCAAGATGC 1140
1141 CCTACACACATGTCATTAAGAGGCACTAGGCTTACACCGGTGCCAGGATTTG 1200
1141 CCTACACACATGTCATTAAGAGGCACTAGGCTTACACCGGTGCCAGGATTTG 1200
1201 GGAAGAGCTGAGCACTCCGTCACCTTCATGAGGAGCTCTTGGCCAAAGTATCA 1260
1201 GGAAGAGCTGAGCACTCCGTCACCTTCATGAGGAGCTCTTGGCCAAAGTATCA 1260
1261 TGTCTCTCTCTCAATTAAGGCTTCAACACCAACCAAAAGTGTGGCCCAAGAGG 1320
1261 TGTCTCTCTCTCAATTAAGGCTTCAACACCAACCAAAAGTGTGGCCCAAGAGG 1320
1321 TGTGTAACCTTTCCGTTTGAACCGGATTTGCTCAACAGCCACGTTTCTGCTCT 1380
1321 TGTGTAACCTTTCCGTTTGAACCGGATTTGCTCAACAGCCACGTTTCTGCTCT 1380
1381 TCTCAGAGAGATCAAGGATCTGATGAGGAAATTCATGCAAGAGCTGAGAGTGG 1440
1381 TCTCAGAGAGATCAAGGATCTGATGAGGAAATTCATGCAAGAGCTGAGAGTGG 1440
1441 CCAAGGCTCTGACCTGCTCCGCTTGAAGCTGCTGATGATGATGATGATGATGATGAT 1500
1441 CCAAGGCTCTGACCTGCTCCGCTTGAAGCTGCTGATGATGATGATGATGATGATGAT 1500
1501 CCAATGACACCTTGTGTAATTCAAAATGAAATCCAACCTGCTCTGAGAGCTTC 1560
1501 CCAATGACACCTTGTGTAATTCAAAATGAAATCCAACCTGCTCTGAGAGCTTC 1560
1561 CTAACCTTGTGAAAGCAAGAGCAAGCTTGAAGGAGCTTCAACCTGCTGCTGCTCT 1620
1561 CTAACCTTGTGAAAGCAAGAGCAAGCTTGAAGGAGCTTCAACCTGCTGCTGCTCT 1620
1621 GACCCCGCTTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
1621 GACCCCGCTTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
1681 CTTTCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1681 CTTTCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1741 CCAAGGCTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1741 CCAAGGCTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1801 AAAGCTGATGTTGAGGAGAACTGAGGCGAGCTTGTGATGATGATGATGATGATGATGAT 1860
1801 AAAGCTGATGTTGAGGAGAACTGAGGCGAGCTTGTGATGATGATGATGATGATGATGAT 1860
1861 TCTTGAATCATGTCAGAGATCCAGAGGCTTAAACCTTGTGCTTGTGAAACCAAGC 1920
1861 TCTTGAATCATGTCAGAGATCCAGAGGCTTAAACCTTGTGCTTGTGAAACCAAGC 1920
1921 TCTGTGCTGAAGGATGAGAGGCTTAAACCTTGTGCTTGTGAAACCAAGC 1980
1921 TCTGTGCTGAAGGATGAGAGGCTTAAACCTTGTGCTTGTGAAACCAAGC 1980

[illegible]

	RESULT 2		
	AX409491		
LOCUS	AX409491	2576 bp	DNA
DEFINITION	Sequence 2138 from Patent WO0229103.	linear	PAT 14-JUN-2002
ACCESSION	AX409491		
VERSION	AX409491.1		
KEYWORDS	GI:21442196		
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1		
AUTHORS	Alvares,C., Horne,D., Peres-da-Silva,S.		
TITLE	Gene expression profiles in liver cancer		
JOURNAL	Patent: WO 0229103-A 2138 11-APR-2002;		
	GENE LOGIC INC (US)		
FEATURES	Location/Qualifiers		
source	1..2576		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/note="EMBL/GenBank Accession No. L04751"		
BASE COUNT	572 a 793 c 583 g 628 t		
ORIGIN			

Query Match	Score 2576	DB 6	Length 2576
Best Local Similarity	100.0%	Pred. No. 0	
Matches 2576	Conservative 0	Mismatches 0	Indels 0
Gaps 0			

Query Match	Score 2576	DB 6	Length 2576
Best Local Similarity	100.0%	Pred. No. 0	
Matches 2576	Conservative 0	Mismatches 0	Indels 0
Gaps 0			

D	1	GAATTCGACAGAAATTCAGCAGGGGTGCGACAAATGAATGCTCTGTGTCAGGCCCGACGA	60
Q	61	GACTTCCTGGGTGATGTCTCTGGAATCTCTCCAAAGGAGCTCCCTGCTCATTTCTGTTTGC	120
D	61	GACTTCCTGGGTGATGTCTCTGGAATCTCTCCAAAGGAGGCTCCCTGCTCATTTCTGTTTGC	120
Q	121	TGCTGATCAAGGCAAGTTCAGCTCAACCTGACAGGCAAGTGGCTGTCAAAAGCCCTCAGC	180
D	121	TGCTGATCAAGGCAAGTTCAGCTCAACCTGACAGGCAAGTGGCTGTCTAAAGCCCTCAGC	180
Q	181	AGTTCGGGTGCGCTCCCTCCCACTGAGTCTCTTGGGACAAATCCAGAGACTCCAAAGAGCC	240
D	181	AGTTCGGGTGCGCTCCCTCCCACTGAGTCTCTTGGGACAAATCCAGAGACTCCAAAGAGCC	240
Q	241	AGGAGCTACACGGAATTCAGAAATGGGTGAGACATTTCCAAAGTGCCTGTCTCATTTGGC	300
D	241	AGGAGCTACACGGAATTCAGAAATGGGTGAGACATTTCCAAAGTGCCTGTCTCATTTGGC	300
Q	301	TATGGGGAGGGAAGTTCGTGTCCAGCTCAATGACCCCTGACTATGAAAGTGAATTCGTG	360
D	301	TATGGGGAGGGAAGTTCGTGTCCAGCTCAATGACCCCTGACTATGAAAGTGAATTCGTG	360
Q	361	GGAAGATGAGACCCGAAATCCCAATGATCTCTAAGATTCGTGCTCCATGGAATTTGGATCG	420
D	361	GGAAGATGAGACCCGAAATCCCAATGATCTCTAAGATTCGTGCTCCATGGAATTTGGATCG	420
Q	421	GCTTGTCTCTGTGAATGGGCAAGCATGTGTTCCAGATCGACGAGATGCTGACCCAGCCT	480
D	421	GCTTGTCTCTGTGAATGGGCAAGCATGTGTTCCAGATCGACGAGATGCTGACCCAGCCT	480
Q	481	TCCACTATGACATCCCTGAAGCCCTAATGGGGGCTCAATGGAGACTCTGTACGATGATG	540
D	481	TCCACTATGACATCCCTGAAGCCCTAATGGGGGCTCAATGGAGACTCTGTACGATGATG	540
Q	541	TGACAATATGGGAAGAGCTCTTGGCCAGGATTTCCCTCTGAGAGTTTTCACAGACGCT	600
D	541	TGACAATATGGGAAGAGCTCTTGGCCAGGATTTCCCTCTGAGAGTTTTCACAGACGCT	600
Q	601	CTTATGATGACTTGGACATCATCATATGAATGTGCTTCAAGCCATCAGGGCAGCATCCAG	660
D	601	CTTATGATGACTTGGACATCATCATATGAATGTGCTTCAAGCCATCAGGGCAGCATCCAG	660
Q	661	TGGAACAGGAATTCAGTCTCTTACATACAGGCCATTTAGTGAACCTGGTTTTTT	720
D	661	TGGAACAGGAATTCAGTCTCTTACATACAGGCCATTTAGTGAACCTGGTTTTTT	720
Q	721	CCCGTGAAGGAATGCTTTCAACAAGATGACCATCTACAGCTGACCTCTGTGGCC	780
D	721	CCCGTGAAGGAATGCTTTCAACAAGATGACCATCTACAGCTGACCTCTGTGTGGCC	780
Q	781	GCTTGAACACCGGCGCTGCGAGCTGGCCATCAGCACACAGACCAAGTATCCAACTGA	840
D	781	GCTTGAACACCGGCGCTGCGAGCTGGCCATCAGCACACAGACCAAGTATCCAACTGA	840
Q	841	GGAAAGCTCAATACAGAGAGAGGGGAGCTGGAAAGATCAAGAGGAAGGCAATTGG	900
D	841	GGAAAGCTCAATACAGAGAGAGGGGAGCTGGAAAGATCAAGAGGAAGGCAATTGG	900
Q	901	ATTTTCTGGAATTCCTCTCTGTGGCCAAATATGAGAAATGGAGAGATTTGTCAACAAG	960
D	901	ATTTTCTGGAATTCCTCTCTGTGGCCAAATATGAGAAATGGAGAGATTTGTCAACAAG	960
Q	961	ACCTCCGTGTGAGGTGAGACGTTCAATTTTGAAGGGCCACGACACCAAGCCAGTGGGA	1020
D	961	ACCTCCGTGTGAGGTGAGACGTTCAATTTTGAAGGGCCACGACACCAAGCCAGTGGGA	1020
Q	1021	TCTCTGGAATCTCTATGCTCTGTGGCCACACCCCAAGCATCAGAGAGGTGCGGGAGG	1080
D	1021	TCTCTGGAATCTCTATGCTCTGTGGCCACACCCCAAGCATCAGAGAGGTGCGGGAGG	1080
Q	1081	AGATTCACACGCTCTGAGGTGAGATGAGAGCCCTCAATCACTGGAACCACTGGAACGATG	1140
D	1081	AGATTCACACGCTCTGAGGTGAGATGAGAGCCCTCAATCACTGGAACCACTGGAACGATG	1140

1141 CCTACACCACTGTCATTAAGAGGCACTGAGGCTTACCCACCGGTCCAGGACTTG 1200
 1141 CCTACACCACTGTCATTAAGAGGCACTGAGGCTTACCCACCGGTCCAGGACTTG 1200
 1201 GAGAGAGCTGAGCACTCCGCTTCACTTCCGTGATGAGGCTCTCTTCCCAAGATCA 1260
 1201 GAGAGAGCTGAGCACTCCGCTTCACTTCCGTGATGAGGCTCTCTTCCCAAGATCA 1260
 1261 TGGTCTCTCTCTCAATTAATGAGCTTCAACAACCAAAAGTGGCCCAACCAAGAG 1320
 1261 TGGTCTCTCTCTCAATTAATGAGCTTCAACAACCAAAAGTGGCCCAACCAAGAG 1320
 1321 TGGTGAACCTTCCGCTTGTGCAACCGGGTCTGCTCAACAAGCTTTCTGCTCCCT 1380
 1321 TGGTGAACCTTCCGCTTGTGCAACCGGGTCTGCTCAACAAGCTTTCTGCTCCCT 1380
 1381 TCTCAGAGAGATCAAGAACTGCAATGGGAAACAATTTGCAAGAGCTGAAGTGG 1440
 1381 TCTCAGAGAGATCAAGAACTGCAATGGGAAACAATTTGCAAGAGCTGAAGTGG 1440
 1441 CCAAGGCGCTGACCTGCTCCGCTTGAAGCTGCTGATCCACAGGATCCCATCC 1500
 1441 CCAAGGCGCTGACCTGCTCCGCTTGAAGCTGCTGATCCACAGGATCCCATCC 1500
 1501 CCAATGCAAGACTTGTGTAATCCAAAAATGGAATCCACTGCTTCAAGAGGCTCC 1560
 1501 CCAATGCAAGACTTGTGTAATCCAAAAATGGAATCCACTGCTTCAAGAGGCTCC 1560
 1561 CTAACTCTGTGAGAACAGAGACAGCTTTGAGAGGCTTCACCTGCTGCTTCTTCT 1620
 1561 CTAACTCTGTGAGAACAGAGACAGCTTTGAGAGGCTTCACCTGCTGCTTCTTCT 1620
 1621 GACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
 1621 GACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
 1681 CTTTCTTCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 1681 CTTTCTTCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 1741 CCAAGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 1741 CCAAGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 1801 AAAGCTGAGTGTGGAGAGAGCTGAGGCTGAGCTGAGTGTGAGTGTAAAGAG 1860
 1801 AAAGCTGAGTGTGGAGAGAGCTGAGGCTGAGCTGAGTGTGAGTGTAAAGAG 1860
 1861 TCTTGAATCAATGCAAGATCCAGAGCTTAAACCCCTTGTGGCTTTGGAACCAAGC 1920
 1861 TCTTGAATCAATGCAAGATCCAGAGCTTAAACCCCTTGTGGCTTTGGAACCAAGC 1920
 1921 TCTTGAATCAATGCAAGATCCAGAGCTTAAACCCCTTGTGGCTTTGGAACCAAGC 1980
 1921 TCTTGAATCAATGCAAGATCCAGAGCTTAAACCCCTTGTGGCTTTGGAACCAAGC 1980
 1981 CCGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 1981 CCGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 2041 GCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 2041 GCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 2101 CATTATCTCAAGTAAACAGACAGATGCTAAACCTGCAACGCTGTAATGCTGCTTAAT 2160
 2101 CATTATCTCAAGTAAACAGACAGATGCTAAACCTGCAACGCTGTAATGCTGCTTAAT 2160
 2161 GCAACATGCTCTTCAACCAACCCCAATCTCAACAGCTGCTTCTTGTGATGATCA 2220
 2161 GCAACATGCTCTTCAACCAACCCCAATCTCAACAGCTGCTTCTTGTGATGATCA 2220

QY 2221 ATAAATATCTGACTTCCAGAGCTCGGGGCTTCAAGGCTTCAATCTTACCTTTGGCG 2280
 DB 2221 ATAAATATCTGACTTCCAGAGCTCGGGGCTTCAAGGCTTCAATCTTACCTTTGGCG 2280
 QY 2281 CCTGAGACCACTTCTCTCAAACTGTCTTTTCACTGCTTGAAGCTTGGCGACTT 2340
 DB 2281 CCTGAGACCACTTCTCTCAAACTGTCTTTTCACTGCTTGAAGCTTGGCGACTT 2340
 QY 2341 TGTACACCCACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
 DB 2341 TGTACACCCACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
 QY 2401 CCTCCCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 DB 2401 CCTCCCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 QY 2461 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 DB 2461 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 QY 2521 TCAGTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2576
 DB 2521 TCAGTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2576

RESULT 3
 HUMCYTFAOH 2576 bp mRNA linear PRI 31-DEC-1994
 LOCUS Human cytochrome p-450 4A (CYP4A) mRNA, complete cds.
 DEFINITION L04751, GI:181396
 ACCESSION L04751, GI:181396
 VERSION L04751.1, GI:181396
 KEYWORDS cytochrome P450; cytochrome P450 4A; fatty acid omega-hydroxylase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2576)
 AUTHORS Palmer, C.N., Richardson, T.H., Griffin, K.J., Hsu, M.H., Muerhoff, A.S., Clark, J.E. and Johnson, E.F.
 TITLE Characterization of a cDNA encoding a human kidney, cytochrome p-450 4A, fatty acid omega-hydroxylase and the cognate enzyme expressed in *Saccharicola coli*
 JOURNAL Biochim. Biophys. Acta 1172 (1-2), 161-166 (1993)
 MEDLINE 93176801
 PUBMED 7679927
 COMMENT Original source text: Homo sapiens adult kidney cDNA to mRNA.
 FEATURES
 source location/Qualifiers
 1..2576
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="kidney"
 /dev_stage="adult"
 1..2576
 /gene="CYP4A"
 33..1592
 /gene="CYP4A"
 /function="fatty acid omega-hydroxylase"
 /standard_name="CYP4A"
 /codon_start=1
 /evidence=experimental
 /product="cytochrome P450"
 /protein_id="AA58436.1"
 /db_xref="GI:181397"
 /translation="MEVSYLSPEKILGDSVGIIOAAILLILLLIKAVOLYLHROWL
 LKALQPPCPSPSHLREHIOLODDLOELQIQRKVEFPFSAQPMWAGCKRVOLYDP
 DYMVKILGRSDPKSHSRYPLAPWIGLILNGQTFQHRMLTPAFHYDILKRYG
 LMAISYVMDKMBELGODSPLEVFQVSAWTLDTIMKCAFSQGS1QVRNSQSY1
 OAISDLNVLVSRVRAHFQNDTYSLSISAGRTWRACOLAHQHTDVTQURKQLOK
 EGHLEKIKRKHLDPLDILLAKMENSILSDKOLRAVDTFMEGHDYTSGISML
 YALATPKHQRHREBHHSLLGDGASITNMHLDQMPYTWTCIKALRLYPPVPGIGRE
 LSTPVPFDDKSLPKGIMVLSIYGAAHNRKVPNPVPDFRPAQPSAOSHAFLEP
 SGGSRNCTGKQFAMNBLKAVATLTLFELLPDTRILPIARLVLSKNGIHLRLR

BASE COUNT 572 a 793 c 583 g 628 t
ORIGIN LRPBBDKXDL

Query Match 100.0%; Score 2576; DB 9; Length 2576;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGCGAGATTCAGACAGGTGCTGCACCATGATGTTCTGTGTGAGCCCCAGCA 60
DB 1 GAATTCGCGAGATTCAGACAGGTGCTGCACCATGATGTTCTGTGTGAGCCCCAGCA 60
QY 61 GACTCCTGGGTATGTTCTTGAAATTCCTCAAGGGGCTCCCTGCTCATTTGCTTCTGC 120
DB 61 GACTCCTGGGTATGTTCTTGAAATTCCTCAAGGGGCTCCCTGCTCATTTGCTTCTGC 120
QY 121 TGTCTATCAAGGACAGTTCAAGCTCTACCTGCACAGGAGTGGCTGCTCAAGACCTTCAG 180
DB 121 TGTCTATCAAGGACAGTTCAAGCTCTACCTGCACAGGAGTGGCTGCTCAAGACCTTCAG 180
QY 181 AGTTCCTGGTCCCTCCCTCCCACTGGCTCTTGCGGACATCCAGAGCTCCAAAGAGACC 240
DB 181 AGTTCCTGGTCCCTCCCTCCCACTGGCTCTTGCGGACATCCAGAGCTCCAAAGAGACC 240
QY 241 AGGAGCTACAGGATTCAGAAATGGGTGAGACATTTCCAAAGTGGCTCTCATTTGGC 300
DB 241 AGGAGCTACAGGATTCAGAAATGGGTGAGACATTTCCAAAGTGGCTCTCATTTGGC 300
QY 301 TATGGGAGGCGAAAGTGGTGTCCAGCTCATAGACCTGACTATATGAAGGTGATTTGCG 360
DB 301 TATGGGAGGCGAAAGTGGTGTCCAGCTCATAGACCTGACTATATGAAGGTGATTTGCG 360
QY 361 GAGATCAAGACCGGAATCCATGTTTCTTACAGATTTCTGGCTCCATGATTTGGGTAG 420
DB 361 GAGATCAAGACCGGAATCCATGTTTCTTACAGATTTCTGGCTCCATGATTTGGGTAG 420
QY 421 GCTTGTCTCTGTTGAATGGGCGACATGTTTCAAGATGAGAGATGCTGACCCAGCTT 480
DB 421 GCTTGTCTCTGTTGAATGGGCGACATGTTTCAAGATGAGAGATGCTGACCCAGCTT 480
QY 481 TCCATATGACATCCGGAAGCCCTATGTTGGGGCTCATGGGACGCTGTGACAGGTATGC 540
DB 481 TCCATATGACATCCGGAAGCCCTATGTTGGGGCTCATGGGACGCTGTGACAGGTATGC 540
QY 541 TGAAGAAATGGGAGAGTCTCTTGGGCAAGATTTCCCTGAGAGTCTTTTACAGAGTCT 600
DB 541 TGAAGAAATGGGAGAGTCTCTTGGGCAAGATTTCCCTGAGAGTCTTTTACAGAGTCT 600
QY 601 CTTGATGACCTTGACACATCATATGATGTGCTTACAGCAATCAGGGCAGATCCAGG 660
DB 601 CTTGATGACCTTGACACATCATATGATGTGCTTACAGCAATCAGGGCAGATCCAGG 660
QY 661 TGAAGAGAAATTCAGATCTTACATACAGGCAATTAATGATGATGATGATGATTTT 720
DB 661 TGAAGAGAAATTCAGATCTTACATACAGGCAATTAATGATGATGATGATGATTTT 720
QY 721 CCGGTGTGAGAAATGCTTTCACAGATGACACATCTTACAGGCTGACCTGCTGCTG 780
DB 721 CCGGTGTGAGAAATGCTTTCACAGATGACACATCTTACAGGCTGACCTGCTGCTG 780
QY 781 GCTGACACACCGGCTGCGAGCTGGCCATCAGACACAGACCAAGATTCACATGA 840
DB 781 GCTGACACACCGGCTGCGAGCTGGCCATCAGACACAGACCAAGATTCACATGA 840
QY 841 GGAAGGCTCACTTACAGAGAGGGGAGCTGAGAAATCAAGAGAGAGGCAATTTGG 900
DB 841 GGAAGGCTCACTTACAGAGAGGGGAGCTGAGAAATCAAGAGAGAGGCAATTTGG 900
QY 901 ATTTCTGGAATATCTCTTGGCCAAATGGAAGATGGAGCATCTTGTGACAGAG 960
DB 901 ATTTCTGGAATATCTCTTGGCCAAATGGAAGATGGAGCATCTTGTGACAGAG 960
QY 961 ACCTCGTGTGAGGTGACAGCTTCAATGTTTGAAGGCGACACACAGACGAGTGGGA 1020

DB 961 ACCTCGTGTGAGGTGACAGCTTCAATGTTTGAAGGCGACACACAGACGAGTGGGA 1020
QY 1021 TCTCTGATTCCTTATGCTCTGGCCACACACCCAGATCAGAGAGTGGCGGAGG 1080
DB 1021 TCTCTGATTCCTTATGCTCTGGCCACACACCCAGATCAGAGAGTGGCGGAGG 1080
QY 1081 AGATTCAGAGCTCTGGGTATGAGGCTCCATCAGCTGGAACACCTGGAACCATGTC 1140
DB 1081 AGATTCAGAGCTCTGGGTATGAGGCTCCATCAGCTGGAACACCTGGAACCATGTC 1140
QY 1141 CCTACACACCAATGATTTAAGAGGCACTGAGGCTTACCAACCGGTGCGAGGATTTG 1200
DB 1141 CCTACACACCAATGATTTAAGAGGCACTGAGGCTTACCAACCGGTGCGAGGATTTG 1200
QY 1201 GAGAGAGCTCAGACATCCCTGACATCTTCATGATGGGCGCTCTTGCCAAAGGTATCA 1260
DB 1201 GAGAGAGCTCAGACATCCCTGACATCTTCATGATGGGCGCTCTTGCCAAAGGTATCA 1260
QY 1261 TGGTCTCTCTTCATTATAGGCTTACCAACCCAAAGTGTGGCCCAACCCAGAGG 1320
DB 1261 TGGTCTCTCTCTTCATTATAGGCTTACCAACCCAAAGTGTGGCCCAACCCAGAGG 1320
QY 1321 TGTGTAACCTTCCGTTTGAACCGGGTTCGCTCAACAGGCGACGGCTTCTGCGCT 1380
DB 1321 TGTGTAACCTTCCGTTTGAACCGGGTTCGCTCAACAGGCGACGGCTTCTGCGCT 1380
QY 1381 TCTCAGAGAGATCAAGAACTCATTTGGGAAACAATTTGCCATGAACGAGTGAAGTGG 1440
DB 1381 TCTCAGAGAGATCAAGAACTCATTTGGGAAACAATTTGCCATGAACGAGTGAAGTGG 1440
QY 1441 CCAAGGCTCTGACCTGCTCCGCTTGAAGCTGTGCTGATCCACAGGATCCCATCC 1500
DB 1441 CCAAGGCTCTGACCTGCTCCGCTTGAAGCTGTGCTGATCCACAGGATCCCATCC 1500
QY 1501 CCAATGACAGATGTTGTTGAATCCAAATGAATTCACCTGGGCTCAGAGAGGCTCC 1560
DB 1501 CCAATGACAGATGTTGTTGAATCCAAATGAATTCACCTGGGCTCAGAGAGGCTCC 1560
QY 1561 CTAACTTGTGAAGAACAAGACAGCTTTGAGGGCTCCACTGCGCTGCTCTTCT 1620
DB 1561 CTAACTTGTGAAGAACAAGACAGCTTTGAGGGCTCCACTGCGCTGCTCTTCT 1620
QY 1621 GAACCCGCTTGTGTCCTTCTCTGCTGCTGATCTGTTTCTGCTGCTGCTTCC 1680
DB 1621 GAACCCGCTTGTGTCCTTCTCTGCTGCTGATCTGTTTCTGCTGCTGCTTCC 1680
QY 1681 CTTTCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCT 1740
DB 1681 CTTTCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCT 1740
QY 1741 CCAAGGCTCCCTACCTGCTTGTACCTGCTCTTACCTGCTTACCTGCTTGTGAGGA 1800
DB 1741 CCAAGGCTCCCTACCTGCTTGTACCTGCTCTTACCTGCTTGTGAGGA 1800
QY 1801 AAAGCTGATGTTGGGAGAGCTGAGGCGGAGCTTGCATGCTGACATATATGTAAGAG 1860
DB 1801 AAAGCTGATGTTGGGAGAGCTGAGGCGGAGCTTGCATGCTGACATATATGTAAGAG 1860
QY 1861 TCTTGAATCATGTCAGATTCAGAGGCTTAAACCTCTTGTGCTTGAACACCAAGC 1920
DB 1861 TCTTGAATCATGTCAGATTCAGAGGCTTAAACCTCTTGTGCTTGAACACCAAGC 1920
QY 1921 TCTGCTGTAAGGAGGAGGCTTACCTTGAAGCAACATATCTTAAAGCGGGCATTA 1980
DB 1921 TCTGCTGTAAGGAGGAGGCTTACCTTGAAGCAACATATCTTAAAGCGGGCATTA 1980
QY 1981 CCCCTGAGGCTTGTGATGATTCAGAGGCTGCTGCTGATGCTGATGCTGCTGCTG 2040
DB 1981 CCCCTGAGGCTTGTGATGATTCAGAGGCTGCTGCTGATGCTGATGCTGCTGCTG 2040
QY 2041 GCTCTTGTCTCTTGTCTCCAGATCAATGATCTTGAAGTAAAGAACTGCTCTC 2100

Db 2041 GCTCCTTGTCTGCTGCTCTCCAGATCAATTGATCTTGAGTTAAAGAACCTGCTTC 2100
 Qy 2101 CATTAATCTCAAGTACAGAGCAGATGCTAAACCGTCAACAGCTGTAATGCTGCTTAAT 2160
 Db 2101 CATTAATCTCAAGTACAGAGCAGATGCTAAACCGTCAACAGCTGTAATGCTGCTTAAT 2160
 Qy 2161 GCACATGCTGCTTGCACCCACCCCTTCTCAACCACTGCTTCTTCTTGTATACCA 2220
 Db 2161 GCACATGCTGCTTGCACCCACCCCTTCTCAACCACTGCTTCTTGTATACCA 2220
 Qy 2221 ATAAATATCTGACCTTCCAGAGCTGAGGAGCTTCAACAGCTGCATCTTGAAGCTTGGAG 2280
 Db 2221 ATAAATATCTGACCTTCCAGAGCTGAGGAGCTTCAACAGCTGCATCTTGAAGCTTGGAG 2280
 Qy 2281 CCTGAGACCACTTCTCTCTCAACCTGCTTCTTCTCACTGCTTGAAGCTTGGAGCTT 2340
 Db 2281 CCTGAGACCACTTCTCTCTCAACCTGCTTCTTCTCACTGCTTGAAGCTTGGAGCTT 2340
 Qy 2341 TGTCAACCCCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
 Db 2341 TGTCAACCCCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
 Qy 2401 CCTCCCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 Db 2401 CCTCCCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
 Qy 2461 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 Db 2461 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 Qy 2521 TCAGTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2576
 Db 2521 TCAGTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2576
 RESULT 4
 LOCUS 567580 2470 bp mRNA linear PRI 26-FEB-1994
 DEFINITION CYP4A11: fatty acid omega-hydroxylase [human, kidney, mRNA, 2470 nt].
 ACCESSION 567580
 VERSION 567580.1 GI:456995
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2470)
 AUTHORS Imaoka, S., Ogawa, H., Kimura, S. and Gonzalez, F.J.
 TITLE Complete cDNA sequence and cDNA-directed expression of CYP4A11, a fatty acid omega-hydroxylase expressed in human kidney
 JOURNAL DNA Cell Biol. 12 (10), 893-899 (1993)
 MEDLINE 94098889
 PUBMED 8274222
 REMARK Genbank staff at the National Library of Medicine created this entry [NCBI gisdb94 141537] from the original journal article.
 FEATURES
 source
 1. 2470
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1. 2470
 /gene="CYP4A11"
 /note="fatty acid omega-hydroxylase, CYP4A11"
 42. 1601
 /gene="CYP4A11"
 /note="major fatty acid-metabolizing P450; This sequence comes from Fig. 1"
 /codon_start=1
 /product="fatty acid omega-hydroxylase"
 /protein_id="AA829502.1"
 /db_xref="GI:456995"

/translation="MSVSYLSPSRLLGDVSGILQASLILILLILIKAVQLYLRHL
 LKALQFPCEPSPHMLFGHIOELQDELORIQKRVTPSPSACHMLMGKAVRQQLDP
 DYKATVLGSDPSHSSTYFLAWIGVIGLLILGQWTFQRRLTPAFHDLKPTVG
 LMAQSVRWLDKREBELLDQSPLEVOHSLMLDTIMKCAFHSQSIQVDRNSQSYI
 QASIDLANVPSRVRAFHQNDITYSLTSGARHTHACQALHPTQVIOLRXQLOK
 ESELEKIKRHLDFDILLAMENGSILSDQLAEVDTPFMEGIDTPASGISLTL
 VALATPHKQERCREBIRHSLDGDASITNNHLDOMVYTMCIKEALRLYPVIGIBR
 LSTPTTPPGRSLPKQIMWLSYGLHNPKWPNBVPSPSPARPSAKSHAPLPF
 SGGSRNCIGKQFAMBELKVAATLALFELLPDPTNIPILARLVLSKNGIHRLRR
 LNPCEDKQDL"
 BASE COUNT 521 a 777 c 579 g 593 t
 ORIGIN
 Query Match 93.8%; Score 2417.2; DB 9; Length 2470;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2443; Conservative 0; Mismatches 3; Indels 4; Gaps 2;
 Qy 7 CGCAGAGATCCAGCAGATGCTGCAACATGATGCTCTGCTGCTGAGCCCGACAGACTCC 66
 Db 16 CTCAGAGATCCAGCAGATGCTGCAACATGATGCTCTGCTGAGCCCGACAGACTCC 75
 Qy 67 TGGGTATGCTCTGGAATCTCTCCAGCGGCTCCCTGCTCATTTCTGCTGCTGA 126
 Db 76 TGGGTATGCTCTGGAATCTCTCCAGCGGCTCCCTGCTCATTTCTGCTGCTGA 135
 Qy 127 TCAAGGAGTTCAAGCTCTACCTGCAAGGAGTGGCTGCTCAAGCCCTCAGACATTC 186
 Db 136 TCAAGGAGTTCAAGCTCTACCTGCAAGGAGTGGCTGCTCAAGCCCTCAGACATTC 195
 Qy 187 CGTGCCTCCCTCCCACTGGCTCTTGGGAGATCCAGAGCTCCAAAGAGAGC 246
 Db 196 CGTGCCTCCCTCCCACTGGCTCTTGGGAGATCCAGAGCTCCAAAGAGAGC 255
 Qy 247 TACAAGGATTCAGAAATGGGTGAGACATTCCTCAAGTGCCTGCTCATTTGGCTATGG 306
 Db 256 TACAAGGATTCAGAAATGGGTGAGACATTCCTCAAGTGCCTGCTCATTTGGCTATGG 315
 Qy 307 GAGGCAAGTTCGTGCTCAGCTCTATGACCTCTATATGAAAGTATTTGGGAGAT 366
 Db 316 GAGGCAAGTTCGTGCTCAGCTCTATGACCTCTATATGAAAGTATTTGGGAGAT 375
 Qy 367 CAGACCCGAAATCCCATGCTTCTTACAGATTCCTGGCTCATAGGATGGGAGGCTTGC 426
 Db 376 CAGACCCGAAATCCCATGCTTCTTACAGATTCCTGGCTCATAGGATGGGAGGCTTGC 435
 Qy 427 TCTGTGTAATGGGAGACATGCTTCCAGATGACGAGTGTGACCCAGCTTTCAC 486
 Db 436 TCTGTGTAATGGGAGACATGCTTCCAGATGACGAGTGTGACCCAGCTTTCAC 495
 Qy 487 ATGACATCTGGAAGCCCTATGAGGAGCTCATGAGACCTCTGACAGTATGCTGAC 546
 Db 496 ATGACATCTGGAAGCCCTATGAGGAGCTCATGAGACCTCTGACAGTATGCTGAC 555
 Qy 547 AATGGAGAGCTCTTGGGAGAGATTCCTCTGAGAGCTTCAACAGTCTCTTGA 606
 Db 556 AATGGAGAGCTCTTGGGAGAGATTCCTCTGAGAGCTTCAACAGTCTCTTGA 615
 Qy 607 TGAACCTGACACCATCATGAGTGTGCTTCAAGCATCAGGAGAGATCCAGGTGACA 666
 Db 616 TGAACCTGACACCATCATGAGTGTGCTTCAAGCATCAGGAGAGATCCAGGTGACA 675
 Qy 667 GGAATCTCAGTCTCATATCAGGAGCTTATGATGATGATCAACCTGATTTTCCGCTG 726
 Db 676 GGAATCTCAGTCTCATATCAGGAGCTTATGATGATGATCAACCTGATTTTCCGCTG 735
 Qy 727 TGAAGATGCTTTCACAGAAAGACATCTACAGCTGAGCTCTGCTGAGCTGGA 786
 Db 736 TGAAGATGCTTTCACAGAAAGACATCTACAGCTGAGCTCTGCTGAGCTGGA 795
 Qy 787 CACACCGGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 846
 Db 796 CACACCGGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 855

/translation="MSVSLSPSRLLDVSGIIQAASLLILLILLIKAVOLYHROWL
LKAHQFPCPSPHMLFPHIOELQDDSLQIKWVSTFPBACHMLMGVAVQUTDP
DYMKVILGSDPSHSHSYRPLAWIYGLILLNGQWFOHRRLTAPAFHDIKPYG
LMAQSVYVMDLKEBELIGODSPLFVFOHSLMTDITMCKAFHOSI QVDRNSOYI
QASIDANTLVFSVRNAPHONDITISLTSGRWTHRAACOLAHOTHOVLOLRKQOK
BEBLEIKKKRMLDPLDILLIIMENGSILSDKILARVDTMPFBGHDITASISMLT
YALATHPKQOERCREBIHSLGDSATITNHLDMPTTMCIGARLVPVGTGRB
LSTPVPFDPGRSLPKGIWMLSLYGLHNKVPMPNBPVPSFPAQSASHAFILP
SGSRNCSIGKQFAMNBLKVATALLRPELLPPTRIPIPIALVKSQKESYCVSG
SLTLVKTARSFEBLHLPSCLPDRFCPLPVCYPVPCLPTFPSSHLPVAVQASPSLS
HLSPLGTCTSLTLPCTISCMKES"

BASE COUNT 521 a 780 c 577 g 591 t

ORIGIN

Query Match 92.9%; Score 2394; DB 9; Length 2469;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2435; Conservative 0; Mismatches 10; Indels 5; Gaps 3;

QY 7 CGCAGAGATTCACAGAGAGTGTGCAACCATGAGTCTCTGTGTGAGGCCCAAGCAACTCC 66
DB 16 CTCAGAGATTCACAGAGAGTGTGCAACCATGAGTCTCTGTGTGAGGCCCAAGCAACTCC 75
QY 67 TGGGTATGCTCTGGAATCTCCAGCGGCTCCCTGCTCATTTGCTCTGTGCTGA 126
DB 76 TGGGTATGCTCTGGAATCTCCAGCGGCTCCCTGCTCATTTGCTCTGTGCTGA 135
QY 127 TCAAGGCAAGTTCAGTCTACCTGCAAGGAGTGTGCTCAAGACCTTCAGCAAGTTCC 186
DB 136 TCAAGGCAAGTTCAGTCTACCTGCAAGGAGTGTGCTCAAGACCTTCAGCAAGTTCC 195
QY 187 CGTGCCTCTCTCCACTGTGCTTGGGCAATCCAGAGCTTCCAAAGCAAGAGC 246
DB 196 CGTGCCTCTCTCCACTGTGCTTGGGCAATCCAGAGCTTCCAAAGCAAGAGC 255
QY 247 TACAAGGATTCGAATGGGTGGAACATTTCCAGTGTGCTCCATTTGGTATGGG 306
DB 256 TACAAGGATTCGAATGGGTGGAACATTTCCAGTGTGCTCCATTTGGTATGGG 315
QY 307 GAGGCAAGTTCGTGTCCAGCTCTATGACCTGACTATATGAAGTATTTCTGGGAGAT 366
DB 316 GAGGCAAGTTCGTGTCCAGCTCTATGACCTGACTATATGAAGTATTTCTGGGAGAT 375
QY 367 CAGACCCGAATCCATGTTCTCTACAGATTCCTGCTCCAGATTTGGGTAACGCTTGC 426
DB 376 CAGACCCGAATCCATGTTCTCTACAGATTCCTGCTCCAGATTTGGGTAACGCTTGC 435
QY 427 TCTGTGTAATGGGCAACATGTTCCAGATTCGAGATGCTGACCCAGCTTCCACT 486
DB 436 TCTGTGTAATGGGCAACATGTTCCAGATTCGAGATGCTGACCCAGCTTCCACT 495
QY 487 ATGACATCTCGAAGCCCTATGTGGGGCTCATGGGAGCTGTATCGAGTATGTGAGCA 546
DB 496 ATGACATCTCGAAGCCCTATGTGGGGCTCATGGGAGCTGTATCGAGTATGTGAGCA 555
QY 547 AATGGGAAGACTCTTGGCCAGAGATTCCTCTGAGAGTCTTTCAAGCAAGTCTCTTGA 606
DB 556 AATGGGAAGACTCTTGGCCAGAGATTCCTCTGAGAGTCTTTCAAGCAAGTCTCTTGA 615
QY 607 TGAACCTGGAACCATATGATGATGCTTCAAGCCATCAAGGAGCAATCCAGGTGAGCA 666
DB 616 TGAACCTGGAACCATATGATGATGCTTCAAGCCATCAAGGAGCAATCCAGGTGAGCA 675
QY 667 GGAATTCAGTCTACTACATACAGGCAATTAAGTGAAGTGAACCAAGTGTGTTTCCCGTG 726
DB 676 GGAATTCAGTCTACTACATACAGGCAATTAAGTGAAGTGAACCAAGTGTGTTTCCCGTG 735
QY 727 TGAAGATGCTCTTCAACAGATGAACCATTAACGCTGACCTTGTGCTGGCCGCTGGA 786
DB 736 TGAAGATGCTCTTCAACAGATGAACCATTAACGCTGACCTTGTGCTGGCCGCTGGA 795
QY 787 CACACCGGCTGTGCAAGTGTGCTGACCAAGCAAGCAAGTATCCAACTGAGGAAG 846
DB 796 CACACCGGCTGTGCAAGTGTGCTGACCAAGCAAGCAAGTATCCAACTGAGGAAG 855

QY 847 CTCACATCAAGAGAGAGGAGAGCTGAGAAATCAAGAGAAAGAGCAATTTGGATTTTC 906
DB 856 CTCACATCAAGAGAGAGGAGAGAGCTGAGAAATCAAGAGAAAGAGCAATTTGGATTTTC 915
QY 907 TGAATATCCCTCTTGGCCAAATAGAGAAATGGAGCATCTTGTGCAAGAGAGCTCC 966
DB 916 TGAATATCCCTCTTGGCCAAATAGAGAAATGGAGCATCTTGTGCAAGAGAGCTCC 975
QY 967 GTGCTAGAGTGAACAGTTCATATGTTGAGGCGCACGACACAGAGCCAGTGGATCTCT 1026
DB 976 GTGCTAGAGTGAACAGTTCATATGTTGAGGCGCACGACACAGAGCCAGTGGATCTCT 1035
QY 1027 GATTCCTCTATGCTCTGGCCACACACCCCAAGATCAAGAGAGTGTCCGGAGAGATCC 1086
DB 1036 GATTCCTCTATGCTCTGGCCACACACCCCAAGATCAAGAGAGTGTCCGGAGAGATCC 1095
QY 1087 ACAGCCCTCTGGGTGATGAGAGCTTCATACCTGGAAACGACTGGACAGATGCTCTACA 1146
DB 1096 ACAGCCCTCTGGGTGATGAGAGCTTCATACCTGGAAACGACTGGACAGATGCTCTACA 1155
QY 1147 CCAACATGTGCAATTAAGAGAGCACTGAGGCTTACCAACCGGTGCAAGGATTTGGCAGAG 1206
DB 1156 CCAACATGTGCAATTAAGAGAGCACTGAGGCTTACCAACCGGTGCAAGGATTTGGCAGAG 1215
QY 1207 AGCTCAGACTCCCGTCACTCTCTGATGGGCGCTCTTGGCCAAAGTATCATGCTCC 1266
DB 1216 AGCTCAGACTCCCGTCACTCTCTGATGGGCGCTCTTGGCCAAAGTATCATGCTCC 1275
QY 1267 TCTCTTCATTTATGACCTTCAACCAACCAAAAGTGGCCCAACCAAGAGTGTTCG 1326
DB 1276 TCTCTTCATTTATGACCTTCAACCAACCAAAAGTGGCCCAACCAAGAGTGTTCG 1335
QY 1327 ACCCTTCCGTTTGGACCGGTTTGTCTCAACAGCCAGCTTCTGCTCTCTCAG 1386
DB 1336 ACCCTTCCGTTTGGACCGGTTTGTCTCAACAGCCAGCTTCTGCTCTCTCAG 1395
QY 1387 GAGATCAAGAACTGATTTGGGAAACAATTTGTCATGAACGAGTGAAGGTGGCCACGG 1446
DB 1396 GAGATCAAGAACTGATTTGGGAAACAATTTGTCATGAACGAGTGAAGGTGGCCACGG 1455
QY 1447 CCTGACCTGCTCCGCTTGAAGCTGTGCTGATCCCAAGAGTCCCATCCCAATTG 1506
DB 1456 CCTGACCTGCTCCGCTTGAAGCTGTGCTGATCCCAAGAGTCCCATCCCAATTG 1515
QY 1507 CACGACGTGTTGAATCCAAATAGAAATTCACCTGAGCTTCAGAGAGCTCCCTAAC 1566
DB 1516 CACGACGTGTTGAATCCAAATAGAAATTCACCTGAGCTTCAGAGAGCTCCCTAAC 1574
QY 1567 CTGTGAGAGCAAGAGCAAGCTTGAAGGCTTCAACCTGACCTGCTCTGACCC 1626
DB 1575 CTGTGAGAGCAAGAGCAAGCTTGAAGGCTTCAACCTGACCTGCTCTGACCC 1634
QY 1627 CGCTTGTCTCCCTTCTGTGTGCGCAATTCCTGTTTCTGTCTGCCCACTTCCCTTCT 1686
DB 1635 CGCTTGTCTCCCTTCTGTGTGCGCAATTCCTGTTTCTGTCTGCCCACTTCCCTTCT 1694
QY 1687 TCCACCTGCTGCTGCGCCAGAGCTGCGGAGCTTCTGCTCTGACCTTCTCAAGG 1745
DB 1695 TCCACCTGCTGCTGCGCCAGAGCTGCGGAGCTTCTGCTCTGACCTTCTCAAGG 1754
QY 1746 CTCCTCACTGTGTGCTACCTGCTCTTCTCAACCACTGATATCTTGTGGAGAAAAGC 1805
DB 1755 CTCCTCACTGTGTGCTACCTGCTCTTCTCAACCACTGATATCTTGTGGAGAAAAGC 1814
QY 1806 TGAATGTTGGAGAGAGCTGAGGCGAGACTTGAATGTGACATATATGAAGAGTCTTG 1865
DB 1815 TGAATGTTGGAGAGAGCTGAGGCGAGACTTGAATGTGACATATATGAAGAGTCTTG 1874
QY 1866 AATCATGTCAAGATCCAGGAGTCTAAACCCCTTGTGGCTTTGGAACACCAAGCTCTGT 1925
DB 1875 AATCATGTCAAGATCCAGGAGTCTAAACCCCTTGTGGCTTTGGAACACCAAGCTCTGT 1934

QY 798 TCCACAGTGGCCATCAGACACACAGCAAGATATCCAACTGAGAAAGCTCAACTACAG 857
 DB 781 TCCACAGTGGCCATCAGACACACAGCAAGATATCCAACTGAGAAAGCTCAACTACAG 840
 QY 858 AAGAGAGGGGAGCTGAGAGAGATCAAGAGAGAGAGCAATTTGATTTTCTGAGATATCTC 917
 DB 841 AAGAGAGGGGAGCTGAGAGAGATCAAGAGAGAGAGCAATTTGATTTTCTGAGATATCTC 900
 QY 918 CTCTTGACCAAAATGAGAGATGAGAGCAATTTGATCAAGAGAGCACTCCGTGAGAGTG 977
 DB 901 CTCTTGACCAAAATGAGAGATGAGAGCAATTTGATCAAGAGAGCACTCCGTGAGAGTG 960
 QY 978 GACAGCTGATTTTGAAGGCGCACACACACAGCAAGAGATTTCTGAGATCTCTAT 1037
 DB 961 GACAGCTGATTTTGAAGGCGCACACACAGCAAGAGATTTCTGAGATCTCTAT 1020
 QY 1038 GCTTGAGCCACACACCCAGATCAAGAGAGATGCGGAGAGAGATCCAGACCTCTG 1097
 DB 1021 GCTTGAGCCACACACCCAGATCAAGAGAGATGCGGAGAGAGATCCAGACCTCTG 1080
 QY 1098 GGTGATGAGGCTTCATCACTGAGAAACAGCTGAGAGATGCGCTTACACACATGTC 1157
 DB 1081 GGTGATGAGGCTTCATCACTGAGAAACAGCTGAGAGATGCGCTTACACACATGTC 1140
 QY 1158 ATTAAGAGAGCACTGAGGCTTACACACAGCTGAGAGATGAGAGAGCTCAAGACT 1217
 DB 1141 ATTAAGAGAGCACTGAGGCTTACACACAGCTGAGAGATGAGAGAGCTCAAGACT 1200
 QY 1218 CCGGTGACCTTCTGATGAGGCGCTCTTGCCCAAGATATGATGCTCTCTCAT 1277
 DB 1201 CCGGTGACCTTCTGATGAGGCGCTCTTGCCCAAGATATGATGCTCTCTCAT 1260
 QY 1278 TATGAGCTTCAACACACCAAAAGTGGGCCAAGAGATTTGACCTTCCGT 1337
 DB 1261 TATGAGCTTCAACACACCAAAAGTGGGCCAAGAGATTTGACCTTCCGT 1320
 QY 1338 TTGACACGGGATCTGCTCAACACAGCAGCTTCTGCTTCAAGAGATCAAG 1397
 DB 1321 TTGACACGGGATCTGCTCAACACAGCAGCTTCTGCTTCAAGAGATCAAG 1380
 QY 1398 AACTGATGAGAGCAATTTGCAATGAGAGCTGAGAGCTGAGAGCTGAG 1457
 DB 1381 AACTGATGAGAGCAATTTGCAATGAGAGCTGAGAGCTGAGAGCTGAG 1440
 QY 1458 CTCCGCTTGAAGCTGCTGATCCACAGAGATCCCAATCCCAATGAGAGCTG 1517
 DB 1441 CTCCGCTTGAAGCTGCTGATCCACAGAGATCCCAATCCCAATGAGAGCTG 1500
 QY 1518 TTGAAATCCAAAATGAGATCACTGCTGAGAGAGCTCCCTAACCCTTGAGAG 1577
 DB 1501 TTGAAATCCAAAATGAGATCACTGCTGAGAGAGCTCCCTAACCCTTGAGAG 1560
 QY 1578 AAGAGACAGCTTGAAGGCGCTCACTGCGCTGCTCTGAGAGAGCTCTGCT 1637
 DB 1561 AAGAGACAGCTTGAAGGCGCTCACTGCGCTGCTCTGAGAGAGCTCTGCT 1620
 QY 1638 CTTTCTGCTGAGAGATCACTGCTGAGAGAGCTCCCTAACCCTTGAGAG 1697
 DB 1621 CTTTCTGCTGAGAGATCACTGCTGAGAGAGCTCCCTAACCCTTGAGAG 1660
 QY 1698 TGTGATCCCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1757
 DB 1681 TGTGATCCCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 QY 1758 TGTGATCCCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1817
 DB 1741 TGTGATCCCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 QY 1818 GAGAGTGAAGGCGAGCTGAGATGATGAGATGATGATGATGATGATGATGAT 1877
 DB 1801 GAGAGTGAAGGCGAGCTGAGATGATGAGATGATGATGATGATGATGATGAT 1860

QY 1878 GATCCAGGCTTAAACCCCTTGTGCTTGTGAGAACACCAAGCTCTGTGAGAGGTG 1937
 DB 1861 GATCCAGGCTTAAACCCCTTGTGCTTGTGAGAACACCAAGCTCTGTGAGAGGTG 1920
 QY 1938 AAGGCTAACCTGAGAGCAATATCTAAGCCCGGAGATTAACCCCTGCTGAGAT 1997
 DB 1921 AAGGCTAACCTGAGAGCAATATCTAAGCCCGGAGATTAACCCCTGCTGAGAT 1979
 QY 1998 AGAATCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2057
 DB 1980 AGAATCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2039
 QY 2058 CTCCAGAGATCAATGATATCTGAGTAAAGAACCGCTCTCATATCTCAAGTACA 2117
 DB 2040 CTCCAGAGATCAATGATATCTGAGTAAAGAACCGCTCTCATATCTCAAGTACA 2099
 QY 2118 GAGCAGATGCTAAACCGTCAAGCTGCTAAATGCTGCTTAATGACAATGCTTTGCA 2177
 DB 2100 GAGCAGATGCTAAACCGTCAAGCTGCTAAATGCTGCTTAATGACAATGCTTTGCA 2159
 QY 2178 CCAACCCCATCTGACCACTGCTTTCTTTGTTATCAACAATAATATCTGACT 2237
 DB 2160 C--CCCACTCTGACCACTGCTTTCTTTGTTATCAACAATAATATCTGACT 2216
 QY 2238 CCAAGCTGAGGCGCTTCAAGCTGCTCAATGCTTGAAGTGGCCCTGAGACCACTTCT 2297
 DB 2217 CCAAGCTGAGGCGCTTCAAGCTGCTCAATGCTTGAAGTGGCCCTGAGACCACTTCT 2276
 QY 2298 CTCTCAACCTGCTTTCTTCTGCTGCTTGAAGTGGCGAGATTTGTCACCCCAAGACT 2357
 DB 2277 CTCTCAACCTGCTTTCTTCTGCTGCTTGAAGTGGCGAGATTTGTCACCCCAAGACT 2336
 QY 2358 GGTGTTGGTCTGAGACACCCCAATCCCTGAATCTTCACCCAGC 2402
 DB 2337 GGTGTTGGTCTGAGACACCCCAATCCCTGAATCTTCACCCAGC 2381

RESULT 7
 AX336099 1763 bp DNA linear PAT 09-JAN-2002
 LOCUS
 DEFINITION
 AX336099
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Buiakoyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horigan, S., Soppet, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 Patent: WO 0194629-A 6608 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 JOURNAL
 TITLE
 FEATURES
 source
 1. 1763
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 381 a 538 c 430 g 414 t
 ORIGIN
 Query Match 63.1%; Score 1625.8; DB 6; Length 1763;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;
 QY 10 AGAGATCAGCAGGCTGCTGACCAATGATGCTGCTGCTGAGCCCGAGCACTCTGG 69
 DB 1 AGAGATCAGCAGGCTGCTGACCAATGATGCTGCTGCTGAGCCCGAGCACTCTGG 60
 QY 70 GTGATGCTCTGGAATCTCCAGAGGCGCTCCGCTCATCTGCTGCTGCTGATCA 129

Db 61 GTGATGTCTGTGAATCTCTCAAGCGGCTCCCTGCTCATTTCTGTCTGTCTATCA 120
 Qy 130 AGGCACTTCACTCTCACTGCAAGGCAAGTGTCTCAAAAGCTTCCAGCACTTCCCT 189
 Db 121 AGGCACTTCACTCTCACTGCAAGGCAAGTGTCTCAAAAGCTTCCAGCACTTCCCT 180
 Qy 130 GGGCTCCCTCCCACTGCTCTTCTGCGGCACTCAAGAGCTTCCCAAGAGCAAGAGCTTAC 249
 Db 181 GGGCTCCCTCCCACTGCTCTTCTGCGGCACTCAAGAGCTTCCCAAGAGCAAGAGCTTAC 240
 Qy 250 AACGATTCAGAAATGGGTGAGAGCAATTCGCAAGTGTCTGTCTCATTTGGCTATGGGAG 309
 Db 241 AACGATTCAGAAATGGGTGAGAGCAATTCGCAAGTGTCTGTCTCATTTGGCTATGGGAG 300
 Qy 310 GCAAAAGTTGTGTCCAGCTCTATGACCTTGAATTAATGAGGTGATTTGGGAGATCAG 369
 Db 301 GCAAAAGTTGTGTCCAGCTCTATGACCTTGAATTAATGAGGTGATTTGGGAGATCAG 360
 Qy 370 ACCGAAATCCCATGCTTCTTACAGATTTCTGCTCATGAGATTTGGGTACGGCTTCTCC 429
 Db 361 ACCGAAATCCCATGCTTCTTACAGATTTCTGCTCATGAGATTTGGGTACGGCTTCTCC 420
 Qy 430 TGTGAATGGGAGCATGCTTCCAGCATCGACGATGTGACCCGAGCTTCCCACTATG 489
 Db 421 TGTGAATGGGAGCATGCTTCCAGCATCGACGATGTGACCCGAGCTTCCCACTATG 480
 Qy 490 ACATCTGAGGCTTATGTGGGCTCATGAGAGCTCTGTACAGAGTATGCTGAGCAAT 549
 Db 481 ACATCTGAGGCTTATGTGGGCTCATGAGAGCTCTGTACAGAGTATGCTGAGCAAT 540
 Qy 550 GGGAGAGCTCTTGGGCGAGATTTCCCTGTGAGGTCTTTTACAGACGTCTCTTGATGA 609
 Db 541 GGGAGAGCTCTTGGGCGAGATTTCCCTGTGAGGTCTTTTACAGACGTCTCTTGATGA 600
 Qy 610 CCTGGAACCATCATGAGTGTGCTTCAAGCATAGGGGAGAGCAATCCAGGTGAGAGAG 669
 Db 601 CCTGGAACCATCATGAGTGTGCTTCAAGCATAGGGGAGAGCAATCCAGGTGAGAGAG 660
 Qy 670 ATTCTCAGTCTTATCATACAGGCTATGAGCTGAAACAAGCTGTTTTTCCCGTGTGA 729
 Db 661 ATTCTCAGTCTTATCATACAGGCTATGAGCTGAAACAAGCTGTTTTTCCCGTGTGA 720
 Qy 730 GGAATGCTTTTCAACAGATTAACCATTTAAGCTTCTGTGCTGCGCTGAGCAC 789
 Db 721 GGAATGCTTTTCAACAGATTAACCATTTAAGCTTCTGTGCTGCGCTGAGCAC 780
 Qy 790 ACCGGGCTTCCAGCTGTGGCTTCAAGCAC--AGAACAATGATTCGAATGAGAGAG 846
 Db 781 ACCGGGCTTCCAGCTGTGGCTTCAAGCAC--AGAACAATGATTCGAATGAGAGAG 840
 Qy 847 CTCACTACAGAAAGGGGGAGCTGAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 906
 Db 841 CTCACTACAGAAAGGGGGAGCTGAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 Qy 907 TGAATATCTCTCTTGTGGCAAAATGAGAGATGAGAGATCTTGTGAGACAGAGACTCC 966
 Db 901 TGAATATCTCTCTTGTGGCAAAATGAGAGATGAGAGATCTTGTGAGACAGAGACTCC 960
 Qy 967 GTGCTGAGGTGAGACAGTTCATGTTTGAAGGCGACACACACAGGCTGAGATCTCTC 1026
 Db 961 GTGCTGAGGTGAGACAGTTCATGTTTGAAGGCGCGACACACAGGCTGAGATCTCTC 1020
 Qy 1027 GGATCTCTATGCTGTGGCGACACACCCGAGATCAGAGAGAGGTGCGGGAGAGAGATCC 1086
 Db 1021 GGATCTCTATGCTGTGGCGACACACCCGAGATCAGAGAGAGGTGCGGGAGAGAGATCC 1080
 Qy 1087 ACAGGCTCTGTGGGTATGAGGCTTCATCATCTGGAACCACTGAGACAGATGCCCTTACA 1146
 Db 1081 ACAGGCTCTGTGGGTATGAGGCTTCATCATCTGGAACCACTGAGACAGATGCCCTTACA 1140
 Qy 1147 CCAAGATGTCATTAAAGAGGCACTGAGGCTTACCCACCGGTGCGAGGATTTGGCAGAG 1206
 Db 1141 CCAAGATGTCATTAAAGAGGCACTGAGGCTTACCCACCGGTGCGAGGATTTGGCAGAG 1200

Qy 1207 AGCTGACACTCCCGTACACTTCCCTGATGAGGCGCTCTCTTGGCCCAAGATATAGTCTC 1266
 Db 1201 AGCTGACACTCCCGTACACTTCCCTGATGAGGCGCTCTCTTGGCCCAAGATATAGTCTC 1260
 Qy 1267 TCTCTTCCATTATAGGCTTTCACCAACCAACCAAAATGTGTGGCCCAACCCAGAGGTGTTG 1326
 Db 1261 TCTCTTCCATTATAGGCTTTCACCAACCAACCAAAATGTGTGGCCCAACCCAGAGGTGTTG 1320
 Qy 1327 ACCCTTCCGTTTGTGACCGGGTTTGTCTCAACAGCCAGCGCTTCTGCTCCCTCTCAG 1386
 Db 1321 ACCCTTCCGTTTGTGACCGGGTTTGTCTCAACAGCCAGCGCTTCTGCTCCCTCTCAG 1380
 Qy 1387 GAGATCAAGAACTGATTTGGGAAACAAATTTGCAATGAACGAGCTGAGAGTGGCC--A 1443
 Db 1381 GAGATCAAGAACTGATTTGGGAAACAAATTTGCAATGAACGAGCTGAGAGTGGCCAGC 1440
 Qy 1444 CGGCTTGAACCTGTCTCGCTTGTGAGCTGTGCTGATCCACAGATCCCATCCCA 1503
 Db 1441 AGGCTTGAACCTGTCTCGCTTGTGAGCTGTGCTGATCCACAGATCCCATCCCA 1500
 Qy 1504 TTGACAGACTTGTGTAATCCAAATGGAATCCAGACCTGCTCAGAGGCTCCCTA 1563
 Db 1501 TTGACAGACTTGTGTAATCCAAATGGAATCCAGACCTGCTCAGAGGCTCCCTA 1560
 Qy 1564 ACCCTTGTGAAGCAAGAACAGCTTGTGAGGCTTCACTGCTGCTTCTTCTGAC 1623
 Db 1561 ACCCTTGTGAAGCAAGAACAGCTTGTGAGGCTTCACTGCTGCTTCTTCTGAC 1620
 Qy 1624 CCGCGCTTCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1683
 Db 1621 CCGCGCTTCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
 Qy 1684 TCTTC 1688
 Db 1681 TCTTC 1685

RESULT 8
 AX336427 1763 bp DNA linear PAT 09-JAN-2002
 LOCUS AX336427
 DEFINITION Sequence 6936 from Patent WO0194629.
 ACCESSION AX336427
 VERSION AX336427.1 GI:18127146
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.,
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 6936 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source
 1. 1763
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 381 a 538 c 430 g 414 t
 ORIGIN
 Query Match 63.1%; Score 1625.8; DB 6; Length 1763;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;
 Qy 10 AGAGATCCGAGAGTGTGTCACCAAGAGTGTCTGTGCTGAGGCCCGCAGAGACTCTGG 69
 Db 1 AGAGATCCGAGAGTGTGTCACCAAGAGTGTCTGTGCTGAGGCCCGCAGAGACTCTGG 60
 Qy 70 GTGATGTCTGTGAATCTCTCAAGCGGCTCCCTGCTCATTTCTGTCTGTCTATCA 129

Db 61 GGAATGCTCTGAAATCTTCAGAGGCGCTCCCTGCTCATTCCTGCTGCTATCA 120
 Qy 130 AGGCAATTCAGCTCTACCTGCAAGGAGTGTGCTCAAGAGCTTCAGAGTTCCCT 189
 Db 121 AGGCAATTCAGCTCTACCTGCAAGGAGTGTGCTCAAGAGCTTCAGAGTTCCCT 180
 Qy 190 GCTCTCTCTCCCACTGCTCTCTCTGAGGCAATCCAGAGCTCCAGAGAGCTTAC 249
 Db 181 GCTCTCTCTCCCACTGCTCTCTCTGAGGCAATCCAGAGCTCCAGAGAGCTTAC 240
 Qy 250 AACGATTCAGAAATGGGTGGAGACATTCGCAAGTCTCTCTCATTTGGCTATG 309
 Db 241 AACGATTCAGAAATGGGTGGAGACATTCGCAAGTCTCTCTCATTTGGCTATG 300
 Qy 310 GCAAGTTGCTGTCTGAGCTCTATGACCTGCTATATGAAAGTGTATTCGGGAGATCAG 369
 Db 301 GCAAGTTGCTGTCTGAGCTCTATGACCTGCTATATGAAAGTGTATTCGGGAGATCAG 360
 Qy 370 ACCGGAATCCAGATGCTCTACAGATTCCTGCTCATGATTTGGGTACGGCTTCTCC 429
 Db 361 ACCGGAATCCAGATGCTCTACAGATTCCTGCTCATGATTTGGGTACGGCTTCTCC 420
 Qy 430 TGTGATTTGGGAGACATGCTTTCAGCATGACAGATTCCTGACCTCCACTTACATATG 489
 Db 421 TGTGATTTGGGAGACATGCTTTCAGCATGACAGATTCCTGACCTCCACTTACATATG 480
 Qy 490 ACATCTGAAGCTCTATGTTGGGCTCATGAGGACATCTGTAAGAGATGCTGGAAGAT 549
 Db 481 ACATCTGAAGCTCTATGTTGGGCTCATGAGGACATCTGTAAGAGATGCTGGAAGAT 540
 Qy 550 GGAAGAGCTCTTGGGCAAGATTCCTCTGAGAGTCTTTCAGACGTCCTCTTATGA 609
 Db 541 GGAAGAGCTCTTGGGCAAGATTCCTCTGAGAGTCTTTCAGACGTCCTCTTATGA 600
 Qy 610 CCTGGAACCATCATGAAATGCTCTTACGACCATGAGGACAGATTCAGGTGACAGGA 669
 Db 601 CCTGGAACCATCATGAAATGCTCTTACGACCATGAGGACAGATTCAGGTGACAGGA 660
 Qy 670 ATTTCAGTCTCATCATAGGAGGCAATTAAGTACCTGAAACAACTGGTTTTTCCCGTGA 729
 Db 661 ATTTCAGTCTCATCATAGGAGGCAATTAAGTACCTGAAACAACTGGTTTTTCCCGTGA 720
 Qy 730 GGAATGCTCTTTCAGAGAAATGACACCATCTACAGCTTACCTGCTGAGCGCTGACAC 789
 Db 721 GGAATGCTCTTTCAGAGAAATGACACCATCTACAGCTTACCTGCTGAGCGCTGACAC 780
 Qy 790 ACCGCGCTGCGAGCTGCGCCATCAGACAC--AGACCAAGTATCAACTGAGAGAG 846
 Db 781 ACCGCGCTGCGAGCTGCGCCATCAGACAC--AGACCAAGTATCAACTGAGAGAG 840
 Qy 847 CTCGAATTCAGAAAGAGGAGGAGCTGAGAAATCAAGAGAGAGGCAATTTGATTTTC 906
 Db 841 CTCGAATTCAGAAAGAGGAGGAGCTGAGAAATCAAGAGAGAGGCAATTTGATTTTC 900
 Qy 907 TGGATATCTCTCTTGGGCAAAATGAGAAATGGAGACATTTGTGACAGAGAGCTTCC 966
 Db 901 TGGATATCTCTCTTGGGCAAAATGAGAAATGGAGACATTTGTGACAGAGAGCTTCC 960
 Qy 967 GTGCTGAGGTGACAGCTTCAATGTTTGAAGGCGCACAGACACAGCGAGTGGATCTCT 1026
 Db 961 GTGCTGAGGTGACAGCTTCAATGTTTGAAGGCGCACAGACACAGCGAGTGGATCTCT 1020
 Qy 1027 GGAATCTCTATGCTCTGCGCACACACCCAGAGATCAGAGAGGTGCGGAGAGATTC 1086
 Db 1021 GGAATCTCTATGCTCTGCGCACACACCCAGAGATCAGAGAGGTGCGGAGAGATTC 1080
 Qy 1087 ACAGCTCTCTGAGTGTGAGCTCTCATCACTGGAACCACTGGAACAGATGCTTACA 1146
 Db 1081 ACAGCTCTCTGAGTGTGAGCTCTCATCACTGGAACCACTGGAACAGATGCTTACA 1140
 Qy 1147 CCAAGATGTCATTAAGAGAGCACTAGAGCTTACCAAGCGGTGCGAGGATTTGGCAGAG 1206

Db 1141 CCAAGATGTCATTAAGAGAGCACTAGAGCTTACCCAGCGGTGCCAGAGATTTGGCAGAG 1200
 Qy 1207 AGCTGAGCACTCCCGTACCTTTCCTGATGAGGCGCTCTTGGCCAAAGTATCATGTCTC 1266
 Db 1201 AGCTGAGCACTCCCGTACCTTTCCTGATGAGGCGCTCTTGGCCAAAGTATCATGTCTC 1260
 Qy 1267 TCTCTCCATTTATAGGCTTTCACCAACCAACCAAAAGTGGGCCCAACCCAGAGGTGTTG 1326
 Db 1261 TCTCTCCATTTATAGGCTTTCACCAACCAACCAAAAGTGGGCCCAACCCAGAGGTGTTG 1320
 Qy 1327 ACCCTTTCCTTTTGGACCGGCTTTCCTGCTACACAGCAAGCTTTCCTGCTCAG 1386
 Db 1321 ACCCTTTCCTTTTGGACCGGCTTTCCTGCTACACAGCAAGCTTTCCTGCTCAG 1380
 Qy 1387 GAGATCAAGAACTGCTGATTTGGAAAACAATTTGCCATGAAAGAGTGGCC--A 1443
 Db 1381 GAGATCAAGAACTGCTGATTTGGAAAACAATTTGCCATGAAAGAGTGGCC--A 1440
 Qy 1444 CGGCTCTGACCTCTGCTCCGCTTTGAGCTGCTGCTGATCCACAGGATCCCATGCCA 1503
 Db 1441 AGGCTCTGACCTCTGCTCCGCTTTGAGCTGCTGCTGATCCACAGGATCCCATGCCA 1500
 Qy 1504 TTGACGACTTGTGTTGAATCCAAATGAAATGCACTGCTGCTCAGAGAGCTCCCTA 1563
 Db 1501 TTGACGACTTGTGTTGAATCCAAATGAAATGCACTGCTGCTCAGAGAGCTCCCTA 1560
 Qy 1564 ACCCTTGTGAAGCAAGAACAGCTTTGAGAGGCTTCACCTGCTGCTGCTTCTGAC 1623
 Db 1561 ACCCTTGTGAAGCAAGAACAGCTTTGAGAGGCTTCACCTGCTGCTGCTTCTGAC 1620
 Qy 1624 CCGCGCTTCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683
 Db 1621 CCGCGCTTCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
 Qy 1684 TCTTC 1688
 Db 1681 TCTTC 1685

RESULT 9
 AX336709
 LOCUS AX336709 1763 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 7218 from Patent WO0194629.
 ACCESSION AX336709
 VERSION AX336709.1 GI:18127428
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Homo sapiens Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCES
 1. Young, P.B., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 Patent: WO 0194629-A 7218 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source
 1. 1763
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 381 a 538 c 430 g 414 t
 ORIGIN
 Query Match 63.1%; Score 1625.8; DB 6; Length 1763;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;
 Qy 10 AGAGATTCAGCAGATGCTGACCATGATGCTCTGCTGAGGCGCCAGAGACTCTGCG 69
 Db 1 AGAGATTCAGCAGATGCTGACCATGATGCTCTGCTGAGGCGCCAGAGACTCTGCG 60

70 GTGATGCTCTGGAATCTCTCAAGCGGCTCTCCCTGCTCATTCGCTTCTGCTGATCA 129
 61 GTGATGCTCTGGAATCTCTCAAGCGGCTCTCCCTGCTCATTCGCTTCTGCTGATCA 120
 130 AGGAGCTTCAAGCTCTCACTGCA CAGGCAATGCTGCTCAAGGCTCTCAAGCTTCCGT 189
 121 AGGAGCTTCAAGCTCTCACTGCA CAGGCAATGCTGCTCAAGGCTCTCAAGCTTCCGT 180
 190 GCGCTCTCTCCCACTGCTCTTCTGCGGCAATCAAGGCTCTCAAGGCTCTCAAGGCTTAC 249
 181 GCGCTCTCTCCCACTGCTCTTCTGCGGCAATCAAGGCTCTCAAGGCTCTCAAGGCTTAC 240
 250 AACGGAATCAGAAATGGGATGAGACATTCGCAAGTCTGCTCTCAATGGGATGAGGAG 309
 241 AACGGAATCAGAAATGGGATGAGACATTCGCAAGTCTGCTCTCAATGGGATGAGGAG 300
 310 GCAAGTTCGCTGCTCAAGCTCTATGACCCCTGCAATATGAGGATGATTCGCGGAGATCAG 369
 301 GCAAGTTCGCTGCTCAAGCTCTATGACCCCTGCAATATGAGGATGATTCGCGGAGATCAG 360
 370 ACCGGAATCCCAATGCTCTCAAGATTCCTGCTCTCAATGATGAGGATGAGGATGCTCC 429
 361 ACCGGAATCCCAATGCTCTCAAGATTCCTGCTCTCAATGATGAGGATGAGGATGCTCC 420
 430 TGTGAAATGGGAGACATGCTTCCAGATTCGAGATGCTGACCCGAGCTTCCATGATG 489
 421 TGTGAAATGGGAGACATGCTTCCAGATTCGAGATGCTGACCCGAGCTTCCATGATG 480
 490 ACATCTGAAAGCTCTATGCTGCGGCTCATGCGAGACTCTGTAAGAGATGCTGAGCAAT 549
 481 ACATCTGAAAGCTCTATGCTGCGGCTCATGCGAGACTCTGTAAGAGATGCTGAGCAAT 540
 550 GGGAAAGACTCTTGGCCAGATTCCTCTGAGGATCTTTCAGCAAGCTCTCTTGATGA 609
 541 GGGAAAGACTCTTGGCCAGATTCCTCTGAGGATCTTTCAGCAAGCTCTCTTGATGA 600
 610 CCGTGAACCATCATGAGATGCTTTCAGCAATCAAGGAGCAATCCAGGATGAGAGAGA 669
 601 CCGTGAACCATCATGAGATGCTTTCAGCAATCAAGGAGCAATCCAGGATGAGAGAGA 660
 670 ATTTCAGTCTCATCATATACAGGAGCATTTAGTGAAGCTGTAACAGCTGTTTTTCCCGTGA 729
 661 ATTTCAGTCTCATCATATACAGGAGCATTTAGTGAAGCTGTAACAGCTGTTTTTCCCGTGA 720
 730 GGAATGCTTTTCAACAGATGACACCATCTACAGCTGCACTCTGCTGCGCTGAGCAAC 789
 721 GGAATGCTTTTCAACAGATGACACCATCTACAGCTGCACTCTGCTGCGCTGAGCAAC 780
 790 ACCGCGCTGCGCAAGCTGCGCTCATGAGCAAC--AGACCAATGATCCATGAGAGAG 846
 781 ACCGCGCTGCGCAAGCTGCGCTCATGAGCAAC--AGACCAATGATCCATGAGAGAG 840
 847 CTCAACTACAGAAAGAGGGGAGCTGAGAGATCAAGAGAGAGAGAGAGAGAGAGAGATTTTC 906
 841 CTCAACTACAGAAAGAGGGGAGCTGAGAGATCAAGAGAGAGAGAGAGAGAGAGATTTTC 900
 907 TGAATATCTCTCTTCTGCGCAAAATGAGAGATGAGAGATCTTCTGCAAGAGAGAGAGAG 966
 901 TGAATATCTCTCTTCTGCGCAAAATGAGAGATGAGAGATCTTCTGCAAGAGAGAGAGAG 960
 967 GTGCTGAGAGTGAACGTTGATGTTGAGAGGCAAGCAACAGAGCAAGCAAGTGGAGATCTCT 1026
 961 GTGCTGAGAGTGAACGTTGATGTTGAGAGGCAAGCAACAGAGCAAGCAAGTGGAGATCTCT 1020
 1027 GGAATCTCTATGCTCTTCTGCGCAACCCCAAGATCAAGAGAGAGAGAGAGAGAGATCTCT 1086
 1021 GGAATCTCTATGCTCTTCTGCGCAACCCCAAGATCAAGAGAGAGAGAGAGAGAGATCTCT 1080
 1087 ACAGGCTCTGAGGATGAGAGCTCATCACTGGAACCACTGGAACCACTGGAACCACTGGAAC 1146
 1081 ACAGGCTCTGAGGATGAGAGCTCATCACTGGAACCACTGGAACCACTGGAACCACTGGAAC 1140
 1147 CCACCATGTGCAATTAAGAGAGCACTGAGGCTCTACCAACCGGTGCGAGGCAATTTGCGAGAG 1206

1141 CCACCATGTGCAATTAAGAGAGCACTGAGGCTCTACCAACCGGTGCGAGGATTTGGCAGAG 1200
 1207 AGCTAGACATCCCGCTCACTTCCCTGATGAGGAGCTCTTTGCCAAGATGATGATGCTCC 1266
 1201 AGCTAGACATCCCGCTCACTTCCCTGATGAGGAGCTCTTTGCCAAGATGATGATGCTCC 1260
 1267 TCTCTTCATTTATGAGCTTTCACCAACCCCAAAATGATGAGGAGAGAGAGAGAGAGAG 1326
 1261 TCTCTTCATTTATGAGCTTTCACCAACCCCAAAATGATGAGGAGAGAGAGAGAGAGAG 1320
 1327 ACCCTTTCGTTTTCACCGGCTTTCCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386
 1321 ACCCTTTCGTTTTCACCGGCTTTCCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 1387 GAGATTCAGGAATGATGATGAGGAAACAAATTTGCCATGAGAGAGAGAGAGAGAGAGAG 1443
 1381 GAGATTCAGGAATGATGATGAGGAAACAAATTTGCCATGAGAGAGAGAGAGAGAGAGAG 1440
 1444 CGGCTTCAGCTCTGCTCGCTTTCAGCTGCTGCTGATCCACAGAGATCCCATCCCA 1503
 1441 AGGCTTCAGCTCTGCTCGCTTTCAGCTGCTGCTGATCCACAGAGATCCCATCCCA 1500
 1504 TTGACAGACTGATGTTGAATCAAAATGAAATGCAACCGGCTTCAGAGAGAGAGAGAG 1563
 1501 TTGACAGACTGATGTTGAATCAAAATGAAATGCAACCGGCTTCAGAGAGAGAGAGAG 1560
 1564 ACCCTTTCAGGAATGATGATGAGGAAACAAATTTGCCATGAGAGAGAGAGAGAGAGAG 1623
 1561 ACCCTTTCAGGAATGATGATGAGGAAACAAATTTGCCATGAGAGAGAGAGAGAGAGAG 1620
 1624 CCGGCTTTCGTTTTCACCGGCTTTCCTGATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1683
 1621 CCGGCTTTCGTTTTCACCGGCTTTCCTGATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1680
 1684 TCTTC 1688
 1681 TCTTC 1685

RESULT 10
 AX408963
 LOCUS AX408963 1763 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence
 ACCESSION AX408963
 VERSION AX408963 1 GI:21441668
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Alvarez, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
 TITLE Gene expression profiles in liver cancer
 JOURNAL Patent: WO 0229103-A 1610 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES
 source Location/Qualifiers
 1..1763
 /organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9606"
 /note="EMBL/GenBank Accession No. D13705"
 BASE COUNT 381 a 538 c 430 g 414 t
 ORIGIN
 Query Match 63.1%; Score 1625.8; DB 6; Length 1763;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;
 10 AGAGATGCAAG 69
 1 AGAGATGCAAG 60


```

OY 70 GTGATGTCCTGGAATCTCCAGCGGCTCCCTGCTCATTCCTGCTTCTGCTGATCA 129
DB 61 GGTATGTCCTGGAATCTCCAGCGGCTCCCTGCTCATTCCTGCTTCTGCTGATCA 120
OY 130 AGGCAATTCAGCTCTAAGCTCTGCAAGGAGGCTGCTCAAGGCTCTCAAGCTTCCGT 189
DB 121 AGGCAATTCAGCTCTAAGCTCTGCAAGGAGGCTGCTCAAGGCTCTCAAGCTTCCGT 180
OY 190 GCGCTCTCTCCCACTGCTGCTCTTGGGCAATCCAGAGCTTCAAGAGCAGAGCTAC 249
DB 181 GCGCTCTCTCCCACTGCTGCTCTTGGGCAATCCAGAGGCTCTCAAGAGCAGAGCTAC 240
OY 250 AACGATTCAGAAATGGGTGGAGACATTCCTCAAGTGTCTCTCATTTGGCTATGGGAG 309
DB 241 AACGATTCAGAAATGGGTGGAGACATTCCTCAAGTGTCTCTCATTTGGCTATGGGAG 300
OY 310 GCAAGTGTCTGTCAGCTCTATGACCTGATCTATATGAGTGTATCTGAGGAGATCAG 369
DB 301 GCAAGTGTCTGTCAGCTCTATGACCTGATCTATATGAGTGTATCTGAGGAGATCAG 360
OY 370 ACCCGAATTCATGCTTCTTACAGATTCCTGCTCTCATGATGATGATGAGCTTGTCTC 429
DB 361 ACCCGAATTCATGCTTCTTACAGATTCCTGCTCTCATGATGATGATGAGCTTGTCTC 420
OY 430 TGTGAAATGGGCAAGATGCTTCAAGATTCAGATGCTGATCCAGCTTCTCACTATG 489
DB 421 TGTGAAATGGGCAAGATGCTTCAAGATTCAGATGCTGATCCAGCTTCTCACTATG 480
OY 490 AACTCTGAAAGCCCTATGTTGGGGCTGATGGGCACTCTGATCGAGATGCTGAGCAAT 549
DB 481 AACTCTGAAAGCCCTATGTTGGGGCTGATGGGCACTCTGATCGAGATGCTGAGCAAT 540
OY 550 GGAAGAAGCTCTCTGAGCAGATTCCTCTGAGAGCTTTCAGACAGTCTCTTGAATGA 609
DB 541 GGAAGAAGCTCTCTGAGCAGATTCCTCTGAGAGCTTTCAGACAGTCTCTTGAATGA 600
OY 610 CCTGTGACACCATGATGATGCTGCTTCAAGCATCAAGGAGAGCTCAAGTGGACAGA 669
DB 601 CCTGTGACACCATGATGATGCTGCTTCAAGCATCAAGGAGAGCTCAAGTGGACAGA 660
OY 670 ATTCTAGTCCTTATCATACAGGCAATTAAGTACCTTGAACCTGCTTCTCCGCTGTA 729
DB 661 ATTCTAGTCCTTATCATACAGGCAATTAAGTACCTTGAACCTGCTTCTCCGCTGTA 720
OY 730 GGAATGCTTTCACAGATGACACCATTAAGCTTCAAGCTTCTGCTGAGCTGAGCAC 789
DB 721 GGAATGCTTTCACAGATGACACCATTAAGCTTCAAGCTTCTGCTGAGCTGAGCAC 780
OY 790 ACCGCGCTGCTGAGCTGAGCTCCATCAAGACAC--AGAACAAGTATCACTAGAGAGG 846
DB 781 ACCGCGCTGCTGAGCTGAGCTCCATCAAGACACAGAGACCAAGTATCACTAGAGAGG 840
OY 847 CTCAACTTACAGAGAGGAGGAGCTGAGAGATCAAGAGAGAGGAGCATTTGATTTTC 906
DB 841 CTCAACTTACAGAGAGGAGGAGGAGCTGAGAGATCAAGAGAGAGGAGCATTTGATTTTC 900
OY 907 TGAATATCTCTCTTGGCCAAAATGAGATGGAGACCTTGTGAGACAAAGACCTCC 966
DB 901 TGAATATCTCTCTTGGCCAAAATGAGATGGAGACCTTGTGAGACAAAGACCTCC 960
OY 967 GTGCTGAGGTGACACGTTATGTTTGAAGGCGCAAGACACACAGCAGTGGATCTCT 1026
DB 961 GTGCTGAGGTGACACGTTATGTTTGAAGGCGCAAGACACACAGCAGTGGATCTCT 1020
OY 1027 GGAATCTCTATGCTCTGAGCACACCCCAAGCATCAAGAGAGTGTCCGGAGAGATCC 1086
DB 1021 GGAATCTCTATGCTCTGAGCACACCCCAAGCATCAAGAGAGTGTCCGGAGAGATCC 1080
OY 1087 ACAGCCCTCCGGGAGTGAAGGCTCCATCACTGGAACCACTTGGACCAAGTGCCTTACA 1146
DB 1081 ACAGCCCTCCGGGAGTGAAGGCTCCATCACTGGAACCACTTGGACCAAGTGCCTTACA 1140
OY 1147 CCAACATGTGCAATTAAGAGGCACTGAGGCTTACCAACCGGTGCCAGGCAATGGCAGAG 1206

```

```

DB 1141 CCAACATGTGCAATTAAGAGGCACTGAGGCTTACCAACCGGTGCCAGGCAATGGCAGAG 1200
OY 1207 AGCTAGACATCCCGTCAACCTTCCATGATGGAGGCTCTTGGCCAAAGATATCATGCTCC 1266
DB 1201 AGCTAGACATCCCGTCAACCTTCCATGATGGAGGCTCTTGGCCAAAGATATCATGCTCC 1260
OY 1267 TCTCTTCATTTATGAGCTTTCACCAACCAAAAGTGTGGCCCAACCCAGAGGTGTTTG 1326
DB 1261 TCTCTTCATTTATGAGCTTTCACCAACCAAAAGTGTGGCCCAACCCAGAGGTGTTTG 1320
OY 1327 ACCCTTCCGTTTGGACCGGCTTGTGCTCAACAGCAGCAGCTTCTGCTCCCTTCTCAG 1386
DB 1321 ACCCTTCCGTTTGGACCGGCTTGTGCTCAACAGCAGCAGCTTCTGCTCCCTTCTCAG 1380
OY 1387 GAGATCAAGGAATGATGATGAGAAACAATTTGCAATGAAAGAGCTGAGGCTGAGC--A 1443
DB 1381 GAGATCAAGGAATGATGATGAGAAACAATTTGCAATGAAAGAGCTGAGGCTGAGCAGC 1440
OY 1444 CGGCTCTGACCTGCTGCTGCTTGAAGTGTGCTGCTGATCCCAACAGATCCCAATCCCA 1503
DB 1441 AGGCTCTGACCTGCTGCTGCTTGAAGTGTGCTGCTGATCCCAACAGATCCCAATCCCA 1500
OY 1504 TTGACAGACTTGTGTTGAATCCAAAATGAAATCCACCTGCTCAGAGAGGCTCCCTA 1563
DB 1501 TTGACAGACTTGTGTTGAATCCAAAATGAAATCCACCTGCTCAGAGAGGCTCCCTA 1560
OY 1564 ACCCTTGTGAAAGCAAGACCAAGCTTGTGAGGCTCCACCTGCTGCTGCTGCTGAC 1623
DB 1561 ACCCTTGTGAAAGCAAGACCAAGCTTGTGAGGCTCCACCTGCTGCTGCTGCTGAC 1620
OY 1624 CCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683
DB 1621 CCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
OY 1684 TCTTC 1688
DB 1681 TCTTC 1685

```

```

RESULT 11
HUMONY
LOCUS
DEFINITION: Human mRNA for fatty acids omega-hydroxylase (cytochrome P-450HKV), complete cde.
ACCESSION D13705
VERSION D13705.1 GI:1197806
KEYWORDS fatty acid omega-hydroxylase; cytochrome P-450HKV.
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Kikuta,Y.
JOURNAL Unpublished
REFERENCE Kikuta,Y.
AUTHORS Kikuta,Y.
TITLE Direct Submision
JOURNAL Submitted (18-NOV-1992) Yasushi Kikuta, Fukuyama University, Faculty of Engineering, Gakuenchou 1, Fukuyama, Hiroshima 729-0292, Japan (Tel:0849-36-2111, Fax:0849-36-2023)
COMMENT
Submitted (18-NOV-1992) to DDBJ by:
Yasushi Kikuta
Department of Food Science and Technology
Fukuyama University
Faculty of Engineering
1 Gakuenchou
Fukuyama, Hiroshima 729-02
Japan
Phone: 0849-36-2111
Fax: 0849-36-2023.
FEATURES
Location/Qualifiers

```



```

source
1. .1763
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HK-7"
/tissue_type="kidney"
/dev_stage="adult"
24. .1589
/notes="heme binding cysteine at AA458."
/codon_start=1
/product="fatty acid omega-hydroxylase"
/protein_id="BAA02864.1"
/db_xref="GI:535787"
/translation="MSVSVLSPSRLGADVGIQLQASLILLLILKAVOLYLRQML
LKALQOFPSPHMLFGHIOELQDDELQIKMVPSPSACPHMLMGKRVOLYD
DMKYLIGSDPKSHSRYFLAPMIGYGLLNGOTWPOHRMLTPAFYDILKPYVG
LMADSVYLDKMBELLGODSLRVRQHSYLMTDITMKCASSHQSIOVDNSQSYI
QASIDNNLVSRVRNRFHNDITYSLSISAGKTHRACQLSHQTRPKYIQLKQALQ
KEGLEKIKRKHLDPLDILLAKWENGSILSDKILRAVDVFMERGHDTYASGLSMI
LYALATPHQRCRBBIHSLGDSISITWNHLDQMPYTMCKEALRLYPPVPGIGR
ELSTVTPDGRSLPGIMVLSIYGLHNPKWMPPEVDFPRFAPGASHALP
PSGSRNCIGKQFAMNELKVGQALTLVAFLLPDPTRIPIPARLVLSKNGIHLRL
RLRNPCEBKDOL"

polya_signal
1736. .1741
polya_site
1763
BASE COUNT 381 a 538 c 430 g 414 t
ORIGIN

Query Match 63.1%; Score 1625.8; DB 9; Length 1763;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;

10 AGAGATCCAGAGAGTCTGACCATGATGTCCTGCTGAGAGCCGACAGACTCTG 69
1 AGAGATCCAGAGAGTCTGACCATGATGTCCTGCTGAGAGCCGACAGACTCTG 60
70 GTGATGTCCTGGAATCTCCAAAGCGGCTCCCTGCTCATTCCTTCTGCTGATCA 129
61 GTGATGTCCTGGAATCTCCAAAGCGGCTCCCTGCTCATTCCTTCTGCTGATCA 120
130 AGGAGTTCAGGCTCTACCTGACAGAGAGTGGTCTCAAGAGCCCTCCAGAGTTCCCGT 189
121 AGGAGTTCAGGCTCTACCTGACAGAGAGTGGTCTCAAGAGCCCTCCAGAGTTCCCGT 180
190 GGCCTCCCTCCAGAGTCTTCCGAGAGATCCAGAGAGTCCAAAGAGAGAGTAC 249
181 GGCCTCCCTCCAGAGTCTTCCGAGAGATCCAGAGAGTCCAAAGAGAGAGTAC 240
250 AACGATTCAGAAATGGGTGAGAGCATTCGCAAGTCTGCTCTCAATGGCTATGGGAG 309
241 AACGATTCAGAAATGGGTGAGAGCATTCGCAAGTCTGCTCTCAATGGCTATGGGAG 300
310 GCAAAGTTCGTGCTCAAGTCTATGACCTGACTATATGAAGGTATTCGGGAGATCA 369
301 GCAAAGTTCGTGCTCAAGTCTATGACCTGACTATATGAAGGTATTCGGGAGATCA 360
370 ACCGAAATCCCATGGTTCCTACAGATTCCTGCTCCATGAGATTCGGGAGATTC 429
361 ACCGAAATCCCATGGTTCCTACAGATTCCTGCTCCATGAGATTCGGGAGATTC 420
430 TTTTGAATGGGAGAGATGGTTCAGCATTCAGAGATCCCAAGCTTCACATAG 489
421 TTTTGAATGGGAGAGATGGTTCAGCATTCAGAGATCCCAAGCTTCACATAG 480
490 ACATCTGAAGCTCTATGTTGGGGCTCATGAGAGATCTGTACAGATGATGCTGACAAAT 549
481 ACATCTGAAGCTCTATGTTGGGGCTCATGAGAGATCTGTACAGATGATGCTGACAAAT 540
550 GGGAGAGATCTCTGGGAGAGATTCCTCTGAGAGTCTTTCAAGCATCTCTCTTATGA 609
541 GGGAGAGATCTCTGGGAGAGATTCCTCTGAGAGTCTTTCAAGCATCTCTCTTATGA 600
610 CCTGAGACATCATGAAGTGTGCTTGAAGCATGAGGAGAGATCCAGAGTGAAGAG 669

```

```

DB 601 CCTGAGACACATCATGAAGTGTGCTTCAAGCATGAGGAGACATTCAGAGTGAACGAG 660
QY 670 ATTCTCAGTCTCATATACAGAGGCTATGATGACCTGAACAACTGTGTTTTTCCGTTGA 729
DB 661 ATTCTCAGTCTCATATACAGGCTATGATGACCTGAACAACTGTGTTTTTCCGTTGA 720
QY 730 GGAATGCTTTTCAACAGAAATGACCATGTACAGCTGACCTTGTGCTGACCTGACAC 789
DB 721 GGAATGCTTTTCAACAGAAATGACCATGTACAGCTGACCTTGTGCTGACCTGACAC 780
QY 790 ACCGCTCTGACAGTCTGACCATGACAC--AAGCAAGATTCACATGAGAGAG 846
DB 781 ACCGCTCTGACAGTCTGACCATGACACAGACCAAGATTCACATGAGAGAG 840
QY 847 CTCACCTACAGAGAGAGGAGAGCTGAGAGATCAAGAGAGAGCATTTGATTTTC 906
DB 841 CTCACCTACAGAGAGAGGAGAGCTGAGAGATCAAGAGAGAGCATTTGATTTTC 900
QY 907 TGGATATCTCTCTTGGCCAAATGAGATGAGAGATCTTGTCAAGACAGACCTCC 966
DB 901 TGGATATCTCTCTTGGCCAAATGAGATGAGAGATCTTGTCAAGACAGACCTCC 960
QY 967 GTGCTGAGGTGACACATTCATGTTGAGGGCCAGACACACAGCATGGGATCTCT 1026
DB 961 GTGCTGAGGTGACACATTCATGTTGAGGGCCAGACACACAGCATGGGATCTCT 1020
QY 1027 GGAATCTCTATGCTCTGAGCCACACACCAAGCATGAGAGAGTCCGGAGAGATCC 1086
DB 1021 GGAATCTCTATGCTCTGAGCCACACACCAAGCATGAGAGAGTCCGGAGAGATCC 1080
QY 1087 ACAGCTCTCTGAGTATGAGAGCTTCATGACCTGAGACCACTGAGACAGATGCCCTTACA 1146
DB 1081 ACAGCTCTCTGAGTATGAGAGCTTCATGACCTGAGACCACTGAGACAGATGCCCTTACA 1140
QY 1147 CCAACATGTCATTAAGAGAGGACATGAGCTCTACCACTGGTCCAGGACATTTGGCAGAG 1206
DB 1141 CCAACATGTCATTAAGAGAGGACATGAGCTCTACCACTGGTCCAGGACATTTGGCAGAG 1200
QY 1207 AGCTGAGCATCCCGGACATTCCTGATGGGGGCTCTTGGCCAAAGATCATGGTCC 1266
DB 1201 AGCTGAGCATCCCGGACATTCCTGATGGGGGCTCTTGGCCAAAGATCATGGTCC 1260
QY 1267 TCTCTCTCATTAATGAGCTTTCACACAACTGAGAGAGAGAGAGAGAGAGTGTG 1326
DB 1261 TCTCTCTCATTAATGAGCTTTCACACAACTGAGAGAGAGAGAGAGAGTGTG 1320
QY 1327 ACCCTTTCGTTTGAACCGGGTTCGTCAACACAGCAGCTTTCCTGCTCAG 1386
DB 1321 ACCCTTTCGTTTGAACCGGGTTCGTCAACACAGCAGCTTTCCTGCTCAG 1380
QY 1387 GAGATTCAGGAATTCGATTTGGGAACAATTTGGCATGAAAGAGCTGAAGTGGCC--A 1443
DB 1381 GAGATTCAGGAATTCGATTTGGGAACAATTTGGCATGAAAGAGCTGAAGTGGCCAGC 1440
QY 1444 CGGCTCTGACCTGCTCCGCTTGAAGCTGCTGATCCCAAGAGATCCCATCCCA 1503
DB 1441 CGGCTCTGACCTGCTCCGCTTGAAGCTGCTGATCCCAAGAGATCCCATCCCA 1500
QY 1504 TTGACGACTTGTGTGAATTCAAAATGGAATTCACCTGCTTCAGAGAGCTCTCTTA 1563
DB 1501 TTGACGACTTGTGTGAATTCAAAATGGAATTCACCTGCTTCAGAGAGCTCTCTTA 1560
QY 1564 ACCCTTGAAGAGAGAGACAGCTTTGAGAGGCTCACTGCGCTCTCTTCTGAG 1623
DB 1561 ACCCTTGAAGAGAGAGACAGCTTTGAGAGGCTCACTGCGCTCTCTTCTGAG 1620
QY 1624 CCGCGCTTCTGCTCCCTGCTGCTGCTGACATCTGTTTCTGCTGACCACTTCCCT 1683
DB 1621 CCGCGCTTCTGCTCCCTGCTGCTGCTGACATCTGTTTCTGCTGACCACTGCTG 1680
QY 1684 TCTTC 1688

```


QY 1147 CCACCATGTCATTAAAGAGGCACTGAGGCTCTACCCACCGGTCCAGGCTATTGGCAG 1206
 DB 1146 CCACCATGTCATTAAAGAGGCACTGAGGCTCTACCCACCGGTCCAGGCTATTGGCAG 1205
 QY 1207 AGCTGAGCACTCCCGTCACTTCCCTGATGAGGCGCTCTTGGCCCAAGGATCATGCTCC 1266
 DB 1206 AGCTGAGCACTCCCGTCACTTCCCTGATGAGGCGCTCTTGGCCCAAGGATCATGCTCC 1265
 QY 1267 TCCTCTCATTTATGAGGCTTCAACACAAACCAAAAGTGGCCCAACCGAGAGTCTTG 1326
 DB 1266 TCCTCTCATTTATGAGGCTTCAACACAAACCAAAAGTGGCCCAACCGAGAGTCTTG 1325
 QY 1327 ACCCTTTCGCTTTGACCGGGCTTCTGCTCAACACAGCAACGCTTCTGCTCTCAG 1386
 DB 1326 ACCCTTTCGCTTTGACCGGGCTTCTGCTCAACACAGCAACGCTTCTGCTCTCAG 1385
 QY 1387 GAGATCAAG 1397
 DB 1386 GAGATCAAG 1396

RESULT 13
 RABCP4A6 1790 bp mRNA linear MM 27-APR-1993
 LOCUS RABCP4A6 1790 bp mRNA, complete cds.
 DEFINITION Rabbit cytochrome P4501VA6 (CYP4A6) mRNA, complete cds.
 ACCESSION M28656
 VERSION M28656.1 GI:164976
 KEYWORDS cytochrome P450; lauric acid omega-hydroxylase.
 SOURCE Oryctolagus cuniculus (rabbit)
 ORGANISM Oryctolagus cuniculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 1 (bases 1 to 1790)
 Johnson, E. F., Walker, D. L., Griffin, K. J., Clark, J. E., Okita, R. T.,
 Merhoffer, A. S. and Masters, B. S.
 Cloning and expression of three rabbit kidney cDNAs encoding lauric
 acid omega-hydroxylases
 Biochemistry 29 (4), 873-879 (1990)
 90254128
 2340280

REFERENCE
 AUTHORS Johnson, E. F., Walker, D. L., Griffin, K. J., Clark, J. E., Okita, R. T.,
 Merhoffer, A. S. and Masters, B. S.
 Cloning and expression of three rabbit kidney cDNAs encoding lauric
 acid omega-hydroxylases
 Biochemistry 29 (4), 873-879 (1990)
 90254128
 2340280
 JOURNAL MEDLINE
 PUBMED
 COMMENT Original source text: Rabbit (strain New Zealand White, adult)
 kidney, cDNA to mRNA, clone Kd46.
 Draft entry and computer-readable sequence for [Biochemistry 29,
 873-879 (1989)] kindly submitted
 by E. F. Johnson, 04-OCT-1989.
 FEATURES
 source
 1. 1790
 /organism="Oryctolagus cuniculus"
 /mol_type="mRNA"
 /db_xref="taxon:9986"
 14. 1546
 /note="lauric acid omega-hydroxylase"
 /codon_start=1
 /protein_id="AAA31230.1"
 /db_xref="GI:164976"
 /translation="MSVSALNPTRLRGLSLGLQVAGLILGLILLLKAAQLYHROWL
 LRLAIOPCPPHMLGSRBRONGHRIQWIKWIKPSPACPRILMGRRLALLVDP
 DYMKVILRSPPKAGSTRPLAPRIGYGLILLNGTWRQHRMLTPAHTYDILKTYG
 LMAISVQIMDMEQLVSQDSLEVPQDISLMTDITKCAFSHSGSVQDLSOSYI
 QAVGDLNLLPFSRVNVPAGSDTYRLSPBGRLSRAQLAHHTDRVLOQRKQLQO
 EGLEIKYRKRLDVLVLPKMGSSLSQDRLRAVDYTMFGHDTASGIMIF
 YALATPHQRHREBIQGLIDGASITWELHDOMPYTTCIKBALRYLPVGVGRQ
 LSPVTPDGRSLPGVIVTLISVALHNPVMPNRPVDPPEPPAPGASARSHALP
 SGGPRNCIGKQPMAMBLKVAVALTVRRLLPDPKRVDPQKRLVLKSNGLHMLRK
 LR"

CDS

BASE COUNT 341 a 610 c 479 g 360 t

ORIGIN

Query Match 48.2%; Score 1241.4; DB 4; Length 1790;
 Best Local Similarity 82.5%; Pred. No. 1.1e-303;
 Matches 1459; Conservative 0; Mismatches 306; Indels 4; Gaps 3;

QY 25 GCTGCACATGAGTGTCTGTGCTGAGGCCCCAGACAGACTCTCTGGGTGATGTCTTGAA 84
 DB 6 GCTGCACATGAGTGTCTGTGCTGAGGCCCCAGACAGACTCTCTGGGTGATGTCTTGAA 85
 QY 85 TCTCTCAAGGGGCTCCCTGCTCAATTCTGCTTGTGCTGATCAAGAGGAGTTCAGGCT 144
 DB 66 TCTCTCAAGGGGCTCCCTGCTCAATTCTGCTTGTGCTGATCAAGAGGAGTTCAGGCT 125
 QY 145 ACCGTGACAGGAGTGGTCTGCTCAAGCCCTCCAGAGTTCCTGCTCCCTCCACT 204
 DB 126 ACTGTGACCGCAGTGGTCTGCTCAAGCCCTCCAGAGTTCCTGCTCCCTCCACT 185
 QY 205 GGTCTTGGGGCATCTCAGAGACTCCAAAGACAGAGCAAGAGCTCAACAGATTCAGAAT 264
 DB 186 GGTCTTGGGGCATCTCAGAGACTCCAAAGAGTTCAAATGGCCATGAGTCAAGTATGTGAAT 245
 QY 265 GGTGAGACATTTCCCAAGTGGCTGCTCAATTGGGCTAAGGGAGCAAGTGTGCTCC 324
 DB 246 GGTGAGACATTTCCCAAGTGGCTGCTCAATTGGGCTAAGGGAGCAAGTGTGCTCC 305
 QY 325 AGCTCTATGACCTTGAATATGAGAGTATTCCTGGGAGATGACACCCGAAATCCCATG 384
 DB 306 TGAATGATACCTTGAATATGAGAGTATTCCTGGGAGATGACACCCGAAATCCCATG 365
 QY 385 GTTCTCTCAAGATTCTGCTCCATGAGATTGGGTACGGCTTGTCTCTGTTGAATGGGCAGA 444
 DB 366 GTTCTCTCAAGATTCTGCTCCATGAGATTGGGTATGGGTATGCTCTGCTGAATGGGCAGA 425
 QY 445 CATTGGTTCAGAGATGAGAGGATGTCACCCGAGCTTCCATCTATGATCTGAAGCCCT 504
 DB 426 CATTGGTTCAGAGATGAGAGGATGTCACCCGAGCTTCCATCTATGATCTGAAGCCCT 485
 QY 505 ATGTGGGCTCATGAGGACACTGTGACAGTATGCTGACAAATGAGGAAGCTCTTGG 564
 DB 486 AGTGGGCTCATGAGGACACTGTGACAGTATGCTGACAAATGAGGAAGCTCTTGG 545
 QY 565 GCGAGATTCCTCTGAGAGTCTTTCAGAGCTCTCTGATGACCTTGACACCATCA 624
 DB 546 GCGAGATTCCTCTGAGAGTCTTTCAGAGCTCTCTGATGACCTTGACACCATCA 605
 QY 625 TGAATGTGCTTCAGAGCATCAGGAGAGATCCAGAGTGAACAGAAATTCAGTCCATCA 684
 DB 606 TGAATGTGCTTCAGAGCATCAGGAGAGATCCAGAGTGAACAGAAATTCAGTCCATCA 665
 QY 685 TACAGGCTATGATGACCTGAACAACTGATTTTCCGTGTGAGAGATGCTTTGACC 744
 DB 666 TCAAGGCTTTGGGACCTGAACAACTGATTTTCCGTGTGAGAGATGCTTTGATC 725
 QY 745 AGAATGACCATCTACAGCTGACCTCTGCTGAGCGCTGAGACACCGGCTTGCCAGC 804
 DB 726 AGAATGACCATCTACAGCTGAGCGCTGAGCGCTGAGCGCTTGCCAGCGCTTGCCAGC 785
 QY 805 TGGCCCATGAGACACAGACCAAGTATTCACATGAGAGAGCTCACTACAGAGAG 864
 DB 786 TGGCCCATGAGACACAGACCAAGTATTCACATGAGAGAGCTCACTACAGAGAG 845
 QY 865 GGGAGCTGAGAGAGATCAAGAGAGAGAGGCAATTGATTTTCTGATATCTCTCTTGG 924
 DB 846 GGGAGCTGAGAGAGATCAAGAGAGAGAGGCGCTTGAATCTCTGAGAGCTCTCTCTTGG 905
 QY 925 CCAAAATGAGAGATGAGAGCATCTTGTCAACAAGAGCTCCGTGTGAGTGGACAGT 984
 DB 906 CCAAGATGAG 965
 QY 985 TCAATGTTAAGGAGGACACACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1044
 DB 966 TCAATGTTAAGGAGGACACACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1025
 QY 1045 CCACACACCCCAAGATCAGAGAGTGCAGGAGAGAGAGATCCACAGCTTCTGGGTGATG 1104
 DB 1026 CCACACACCCCAAGATCAGAGAGTGCAGGAGAGAGAGATCCACAGCTTCTGGGTGATG 1085
 QY 1105 GAGCTTCATCACTGGAACCACTGACCAAGATGCTTACACACAGATGTGATTAAG 1164

```

Db      1086 GAGCTCCATCAGCTGGAGACCTGAGACAGAGTCCCTACACACCATGTGATCAAGG 1145
Qy      1165 AGGCACTGAGGCTTACCAACGAGTGCAGAGCTATTGGACAGAGCTCAGACCTCCGCTCA 1224
Db      1146 AGGCGCTGAGACTTACCAACGAGTGCAGAGTGCAGAGCTCAGCTCAGCTCAGCTGTCA 1205
Qy      1225 CTTTCCCTGATGAGGCGCTCTTGGCCCAAGATCATGTCTCTCTCTCTCTCTCTCTCTCT 1284
Db      1206 CTTTCCCTGATGAGGCGCTCTTGGCCCAAGATCATGTCTCTCTCTCTCTCTCTCTCTCT 1265
Qy      1285 TTGACCAACACCAACAGTGTGGCCCAACCAAGTGTGTGACCTCTTCCGTTTGCAC 1344
Db      1266 TTGACCAACACCAACAGTGTGGCCCAACCAAGTGTGTGACCTCTTCCGTTTGCAC 1325
Qy      1345 CGGGTTCTGCTCAACACAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1404
Db      1326 CGGGTTCTGCTCAACACAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1385
Qy      1405 TTGGGAAACAATTGTCATGAAAGAGCTGAAAGTGGCCACGCGCTGACCTCTCTCTCTCT 1464
Db      1386 TCGGGAAGCAATTGTCATGAAAGAGCTGAAAGTGGCCACGCGCTGACCTCTCTCTCTCT 1445
Qy      1465 TTGAGCTGCTGCTGATTCACACAGATGCCCATTCCTCATGACAGATTGTGTGAAT 1524
Db      1446 TCGAGCTGCTGCTGATTCACACAGATGCCCATTCCTCATGACAGATTGTGTGAAT 1505
Qy      1525 CCAAAAATGAAATTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1583
Db      1506 CCAAGCAACGAGATTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1565
Qy      1584 CAGCTTTGAGGGGCTTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1643
Db      1566 AGGCTCTGAGGGGCTTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1625
Qy      1644 TGTCTGCGCAATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1703
Db      1626 TGTCTGCGCAATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1685
Qy      1704 CCCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1761
Db      1686 CTTTGGGCTTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1745
Qy      1762 CTACCTGCTCTCTACCACTGATCTCT 1790
Db      1746 CTA-CGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1773

```

RESULT 14
RABCTYR9 2191 bp mRNA linear MAM 27-Apr-1993
LOCUS RABCTYR9 2191 bp mRNA, complete cds.

DEFINITION Rabbit cytochrome P-450-kal mRNA, complete cds.
ACCESSION M29531.1 GI:164986
VERSION M29531.1 GI:164986
KEYWORDS cytochrome P450kal.
SOURCE Oryctolagus cuniculus (rabbit)
ORGANISM Oryctolagus cuniculus

REFERENCE 1 (bases 1 to 2191)
AUTHORS Yokotani, N., Bernhardt, R., Sogawa, K., Kusunose, E., Gotch, O.,
Kusunose, M., and Fujii-Kuriyama, Y.
TITLE Two forms of omega-hydroxylase toward prostaglandin A and laurate.
J. Biol. Chem. 264 (36), 21665-21669 (1989)

JOURNAL MEDLINE 90094341
PUBMED 2600085
COMMENT Original source text: Rabbit (Japanese white) adult male kidney,
cDNA to mRNA, clone R9.
Draft entry and printed sequence for [1] kindly submitted by
N.Yokotani, 27-Oct-1989.

FEATURES
source 1..2191
Location/Qualifiers

```

CDS
/organism="Oryctolagus cuniculus"
/mol_type="mRNA"
/db_xref="taxon:9986"
10..1542
/note="cytochrome P-450-kal (EC 1.14.99.)"
/codon_start=1
/protein_id="A0A31334.1"
/db_xref="GI:164987"
/translation="MSVSALNPTRLPGSLGLQVAGLLGLLLKRAQLYHRLNL
LRALQPPCPFFHMLGHSRFPQNGHELQVLMKVEKFPSPACPMWLSGSHLLIYD
DYMKVILGRSDPRAQSYRFLAIVIGYGLILNGQTFQHRRLMTPAFHDIAPYG
LMAQSVQIMDKBOLVSDSLFVODLSIMLDTIMKCAFSGHSGVQLDRNSQYI
QAVGDMNLPRSRVVRPHQSDPTVYLSLRGSHSAACOLAHHTHDVIOORXALO
ESBLERVRKRRLDLDVILPAMENSGSSISDDDLAAYDTFMEGHDTPASGISMIF
VALATPHERHCRRESIQGLIDGASITMELHMQMPPTTICIKBALYLPVPQVGR
LSPVTPPGRSLIPKGVITVLSLYALHNHKKCPNPEVDPSPAPGSAHSHAFLEP
SGPRNICQKQFAMNELKVAVALTLRFLPDKVPPDKPRVLVKSNGIHRLRK
LR"
BASE COUNT 407 a 744 c 558 g 482 t
ORIGIN
Query Match 47.3%; Score 1219; DB 4; Length 2191;
Best Local Similarity 81.7%; Pred. No. 5,5e-298;
Matches 1445; Conservative 0; Mismatches 320; Indels 4; Gaps 3;
25 GCTGCACATGAGTGTCTGTCTGAGGCCCAAGACATCTCTGGGTGATGTCTTGAA 84
Db 2 GCTGCACATGAGCGTGTCTGACATGAAACCCGAGCTCCGGGAGCCTCTCCGAGC 61
Qy 85 TCCCTCAAGCGGCTCCCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 144
Db 62 TCCCTCAAGTGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Qy 145 ACCTGCAACAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 204
Db 122 ACCTGCAACCGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
Qy 205 GGGCTTTCGGGACATTCACAGAGCTCCACAGACACAGAGCTTCAACGATTCAGAAAT 264
Db 182 GGGCTTTCGGGACACCGCGAGATTCACAAATGCGCAATGATGATGATGATGAT 241
Qy 265 GGGTGAAGACATTCAGAGCGCTGCTCATTTGGGCTATGAGGAGGAGGAGGAGGAGGAGGAG 324
Db 242 GGGTGAAGAAATTCAGAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
Qy 325 AGCTCATGACCGCTGCTATATAGAGTGTCTGAGGAGATCAGACCCGAAATCCCATG 384
Db 302 TGAATCATGACCGCTGCTATATAGAGTGTCTGAGGAGATCAGACCCGAAATCCCATG 361
Qy 385 GTTCTCAAGATTCCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 444
Db 362 GTTCTCAAGATTCCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
Qy 445 CATGCTTTCAGATTCACAGATGTGACCCCAAGCTTCCATGACATCTCTGAAGCCCT 504
Db 422 CGTGTTCACAGACCGCGCATGCTCACCCAGCTTCCATGACATCTCTGAAGCCCT 481
Qy 505 ATGAGGGGCTCATGACAGCTGTGACAGTGTGATGATGATGATGATGATGATGATGATGATGAT 564
Db 482 ACCTGAGGCTCATGAGCGGATCTCGTCAATCATGCTGACAAATGGAGAGAGCGTCA 541
Qy 565 GCCAGAGATTCCTCTGAGAGTCTTTTCAAGACGTCCTCTTGATGATGATGATGATGATGATGAT 624
Db 542 GCCAGAGATTCCTCTGAGAGTCTTTTCAAGACATCTCTTGATGATGATGATGATGATGATGAT 601
Qy 625 TGAAGTGTGCTTCAAGCATCAGGAGGAGCATTCAGAGTGTGACAGAAATCTCAGTCTTCA 684
Db 602 TGAAGTGTGCTTCAAGCATCAGGAGGAGGCTTCAAGAGAAATCCAGTCTTCA 661
Qy 685 TACAGGCTTGAAGACCTGAAACACTGTTTTCCTGCTGATGAGAAATCTTTCAC 744
Db 662 TCAAGCTGTGAGGAGCTTCAACACTGTTTTCCTGCTGATGAGAAATCTTTCATC 721

```

745 AGAATGACACCATCTACAGCCTGACCTCTGCTGCGCTGAGCAACCGCGCTGCGACG 804
 722 AGAGTGAACACCATCTACAGCCTGAGCCTCTGAAAGCGGCTGTGCTCCACCGTGTCTCGACG 781
 805 TGCGCCATCAGACACAGACCAAGTATCTCAACTGAGAGAGGCTCAACTACAGAGAGG 864
 782 TCGCCACAGACACAGACCAAGTATCTCAAGAGAGAGGCTCAAGTCAAGAGAGG 841
 865 GGGAGCTGAGAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
 842 GGGAGCTGAGAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
 925 CCAAAAG 984
 902 CCAAG 961
 985 TCATGTTGAG 1044
 962 TCATGTTGAG 1021
 1045 CCACACACCCCAAGACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1104
 1022 CCAGACACCCCAAGACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
 1105 GAGGCTGATCACTGAG 1164
 1082 GAGGCTGATCACTGAG 1141
 1165 AGGACATGAGAGCTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1224
 1142 AGGAGCTGAGAGCTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201
 1225 CTTTCCTGATGAG 1284
 1202 CTTTCCTGATGAG 1261
 1285 TTGACCAACACCAAG 1344
 1262 TTGACCAACACCAAG 1321
 1345 CGGCTTCGCTGACACAG 1404
 1322 CGGCTTCGCTGACACAG 1381
 1405 TTGCGAAACCAATTTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1464
 1382 TCGGGAAGCAATTTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1441
 1465 TTGAGCTGCTGCTGATCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1524
 1442 TCGAGCTGCTGCTGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
 1525 CCAAAAG 1583
 1502 CCAGCAACCGGAGATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1561
 1584 CAGCTTTGAG 1643
 1562 AGGCTTCGAG 1621
 1644 TGTCTGCGCAATCTGTTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1703
 1622 TGTCTGCGCAATCTGTTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
 1704 CCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1761
 1682 CTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741
 1762 CTACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1790
 1742 CTA-CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1769

RESULT 15
 LOCUS OCP450KD 2377 bp mRNA linear MAM 22-MAY-1991
 DEFINITION Rabbit mRNA for omega-hydroxylase cytochrome P-450 (kd).
 ACCESSION X57209
 VERSION X57209.1 GI:1655
 KEYWORDS Cytochrome P450.
 SOURCE Oryctolagus cuniculus (rabbit)
 ORGANISM Oryctolagus cuniculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 REFERENCE 1 (bases 1 to 2377)
 Yokotani, N., Kusunose, R., Sogawa, K., Kawashima, H., Kinoshita, M.,
 Kusunose, M., and Fujii-Kuriyama, Y.
 cDNA cloning and expression of the mRNA for cytochrome P-450kd
 which shows a fatty acid omega-hydroxylating activity
 Eur. J. Biochem. 196 (3), 531-536 (1991)
 JOURNAL 91192021
 MEDLINE 2013275
 PUBMED 2 (bases 1 to 2377)
 Fujii-Kuriyama, Y.
 Direct Submission
 Submitted (28-DEC-1990) Y. Fujii-Kuriyama, DEPT OF BIOCHEMISTRY,
 FACULTY OF SCIENCE, TOHOKU UNIVERSITY, ARAMAKI AOB-KU, SENDAI 980,
 JAPAN
 FEATURES
 source
 1..2377
 /organism="Oryctolagus cuniculus"
 /mol_type="mRNA"
 /db_xref="taxon:9986"
 /tissue_type="kidney"
 1..2377
 /gene="P-450 (kd)"
 1..2377
 /gene="P-450 (kd)"
 /evidence=experimental
 22..1557
 /gene="P-450 (kd)"
 /codon_start=1
 /product="omega-hydroxylase cytochrome P-450"
 /protein_id="CAA40493.1"
 /db_xref="gi:1655"
 /db_xref="SWISS-PROT:P14579"
 /translation="MSVSALSPTRPGLSLGLQVALLGLLLIKAAQLYRQWL
 LRALQPPCPFHLIHSRBFQWNOBLOQLIKVEKPRACPHMIGSKVRVQLYDP
 DYNVILGRSDPKSGSYTPVAPWIGQLILNQPWFQHRMLTPAFYDLKPYVG
 LMDSVQIMLKMQLVSDSLRVDQISLMTDTIKCARSYGSLDSDNSQY
 IOAVGDLNLLVFAVRNIFHSDPTIYRLSPERLSHRACQLAHEHTRVIOQRKALQ
 QEGELKVRRLRDLFDVLLPARKENSSLSDDLRVDFTFEGHDTTSGVSWI
 FYALATHEHGRCHREIQLIGLDASITWELDMPTTMCIKEMARIPVPVALP
 DLSPTVEPDGRSLPKGFTVLISYIGLHNPNVNPVEFDSRFPGSARSHALP
 PGGARNICIGKPMNBLKVALVLLRPELLPRTIPKPYARLVLKSNNGIHLRLR
 KLV"
 2342..2347
 /gene="P-450 (kd)"
 polyA_signal
 2342..2347
 /gene="P-450 (kd)"
 BASE COUNT 497 a 758 c 565 g 557 t
 ORIGIN
 Query Match 46.5%; Score 1197.2; DB 4; Length 2377;
 Best Local Similarity 83.4%; Pred. No. 1.9e-292;
 Matches 1396; Conservative 0; Mismatches 273; Indels 5; Gaps 3;
 21 AGGTGCTCACCATAGTGTCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 80
 10 AGCGCTGACACATAGAGGTGTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 69
 81 GGAATCTTCAAG 140
 70 GGGCTCTTCAAG 129
 141 CTCTACCTGACAG 200
 130 CTCTACCTGACAG 189

OY		201	CACGAGCCTTCGGGCACATCCAGGAGCTCCAAAGAGCACAGAGACTCAACGGATTCCAG	260
Db		190	CATCGCTTCTGGGGCACAGCCGAGAGTTCCAGATGAACAAGAGCTTCACACAGATTCTG	249
OY		261	AAATGGGTGAGAGCATTTCCCAAATGCCCTGTCTCTCAATTGGCTATGGGAGGCAAGTTTGGT	320
Db		250	AAATGGGTGAGAGAGTTTCCCAAAGAGCCTGTCTCACTGGATATAGGGGGGCAACAAAGTCCGA	309
OY		321	GTCAGCTCTATGACCCTGACTATATGAAAGTGATTCTGGGGAGATCAGACCCGAAATCC	380
Db		310	GTCAGCTCTATGACCCCGAATACTGAAGGTGATCTGGGGAGATCAGATCCAAAGCTC	369
OY		381	CATGGTTCCTACAGATTCTGGGCTCAATGAAATGGGTAGAGGGCTTGCTCTGTGAATGGG	440
Db		370	CTGGTTCCTACACATTGTGGCTCTCGAATGGGTAGAGGGTTGTCTCTCTGAACGGG	429
OY		441	CAGACATGGTTCCAGCATTCGAGATCGAGATGCTGACCCCAAGCTTTCATATGACATCTTGAA	500
Db		430	CAGGCATGGTTCCAGCACCGGGGCGATGTCAACCCAGCTTTCATACAGACATCTTGAA	489
OY		501	CCCTATGTGGGGCTCATGGCAGACTCTGTACGATGATGCTGGAACAAATGGGAAAGCTC	560
Db		490	CCCTAATGGGGGCTCATGTGTGACTCCGTCMAAATCATGTGGAACAAATGGGAGAGCTG	549
OY		561	CTTGGCCAGGAATCCCCTCTGGAAGTCTTTTCAAGACGTCCTCTTGATGATACCTTGACACC	620
Db		550	GTCAGCCAGGACTCCTCCCTGGAAGTCTTCCAAAGCATCTCCCTGATGACCTCTGGAACCC	609
OY		621	ATCATGAAGTGTGCTTCAGGCATATCAGGGGAGACATCCAGGTGGA---CAGGAATTCTCAG	677
Db		610	ATCATGAAGTGTGCTTCAGGCTTAAGCTAACAGGGCAGGCTTCAGCTGGAACAGAGAAATTCAG	669
OY		678	TCTCATATACAGGCCCATTAATGACCTTGAAACAACCTGTTTTTTTCCCTGTGAGGAATGCC	737
Db		670	TCTCATATCCAGGGGCTGTGGGGACCTGAAACAATTGGTTTGGCCGGGTGAGGAATTC	729
OY		738	TTTCAACCAAGAAATGACACAATCAAGCTTCAAGCTTCTGCTGGCCGCTTGACACACCGGCC	797
Db		730	TTTCAACCAAGACGACACCATCTACAGGCTGAGGCCCTGAAAGGCCGCTTCTCCACCTGTGCC	789
OY		798	TGCCAGCTGGCCCATCAGACACACACAAAGATTCCAATGAGGAAGGCTCAACTACAG	857
Db		790	TGCCAGCTGGCCCATGAGACACACACCGAGATTCAGCAGAGGAAGCTCAGCTGACG	849
OY		858	AAGAGGGGGAGCTGAGAGATCAAGAGGAAGGCAATTTGATTTCTGGAATPACTCTC	917
Db		850	CAGAGGGGGAGCTGAGAGAGTCAAGAGGAAGGCGCTTGGAATTCTTGACACTCTTC	909
OY		918	CTTTGGCCAAAATGAGAATGGAGACATTGTTCAGACAAAGACCTCCGTCGAGGTG	977
Db		910	CTTTTGGCCAAAGATGAGAGAGGAGCAGCTGTCCAGACAGGACCTCCGCGCCGAGGTG	969
OY		978	GACAGCTTCAATGTTTGAAGGGCCAAGACACACAGCAGATGGGATCTCTGGAATCCTCTAT	1033
Db		970	GACAGCTTCAATGTTTGAAGGGCCAAGACACACAGGCAAGCGGGTCTCTGGAATCTTCTAT	1028
OY		1038	GCTCTGGCCACACCCCAAGCATCAGAGAGGATGCCGGGAGAGATCCACAGCTCTCTG	1097
Db		1030	GCTCTGGCCACGACACCCCGAGCATCAGACCGGATGCCGGGAGAGATCCAGGACTCTCTG	1088
OY		1098	GGTGATGAGGCTTCATCACTCTGGAACCACTTGACAGATGCCCTTACACCACTATGTC	1157
Db		1090	GGGAGCGGAGCTTCATCACTCTGGGAGGACCTTGACAGATGCCCTTACACCACTATGTC	1148
OY		1158	ATTAAAGAGGCACTGAGGCTCTAACCAACGGGTGCCAGGCAATGGCAGAGAGCTCAGCACT	1211
Db		1150	ATCAAGAGGCGCATATGACTTACCAACCGGGTCCAGGCTATTAATGATGAGACTTCAGCTCA	1208
OY		1218	CCCGTCACTTCCCTGATGGGGGCTCTTGGCCCAAGGATCATGGTCTCTCTCCATYT	1277
Db		1210	CTGTGACCTTCCCTGATGGAAGGCTCTCTCCCCCAAGGTTTTCAGAGTCAACTCTTCATYT	1266

QY	1278	TATGGCCCTTCAACCAACCAACCAAAAGTGTGGCCCAACCAAGAGGTGTTGACCTTTCGCT	1337
Db	1270	TATGGCCCTTCAACCAACCAACCAAAAGTGTGGCCCAACCAAGAGGTGTTGACCTTTCGCT	1329
QY	1338	TTTGGCACTGGGGTTCCTGTCAACACGACACGCTTTCCTGCCCCCTTTCAGAGAGATCAAGG	1397
Db	1330	TTTCAACAAGGTTCTGTCTGGCAACGACACGCTTTCCTGCCCCCTTTCAGAGAGAGCAAGG	1389
QY	1398	AACCTGCAATTGGGAAACAATTTTGCATGACAGAGCTGAAAGGTGGCCACGGCCCTGACCTTG	1457
Db	1390	AACCTGCAATTGGGAAAGCAATTTGCGCATTAACAGAGCTGAAAGGTGGCCCGTCAACCTTG	1449
QY	1458	CTCCGCTTTGAGCTGCTGCTGATGCCACACAGATATCCCATTCCTCATTTGACAGACTTGTG	1517
Db	1450	CTGGGCTTTCAGACTGCTGCTGCGGATCCACACAGATATCCCATTCCTCATTTGACAGACTTGTG	1509
QY	1518	TTGAAATCCAAAAATGAAATTCACCTGCTCAGAGGCTCC-CTAACCTTGTGAAGA	1576
Db	1510	CTGAAGTCCAAACAAGGGATTCACCTGCTGTGAAGAGCTCCAAATACCTGCTGGGGA	1568
QY	1577	CAAGAACAGCTTTGAGGGCTCC-ACTGCGCTCTGTCTTTCCTGACCCCCGCTTCTGT	1635
Db	1570	CACGAGTAGAACAAGGAGCTCCTCTGAATGCCATCTTTCCTGCCCCATTTTCTCGG	1629
QY	1636	CCCCCTCTGTGCGCCCAATACCGTCTTTCGTGTGCGCAACCTTCCCTTCTTC	1689
Db	1630	GCTTTGCAATCTACCGCATCTGCTTTCATATGGCCACCTTGAAGCTTCTTC	1683

Search completed: February 15, 2004, 03:56:54
Job time : 9395 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 16:07:28 ; Search time 672 Seconds
(without alignments)
10347.846 Million cell updates/sec

Title: US-09-942-429A-6

Perfect score: 2576
Sequence: 1 gaatccgcagagatccagc.....aaaaaaaaaacgcgaatc 2576

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2576	100.0	2576	24	ABN95640
2	2576	100.0	2576	24	ABN95640
3	2576	100.0	2576	24	ABN95640
4	1625.8	63.1	1763	24	ABL68591
5	1625.8	63.1	1763	24	ABL68591
6	1625.8	63.1	1763	24	ABL68591
7	1625.8	63.1	1763	24	ABL68591
8	1525.2	59.2	1872	24	ABK50070

9	1082.8	42.0	2116	24	ABK50068	DNA encoding mouse
10	1022.4	39.7	2452	24	ABK63715	Rat sequence diffe
11	968.4	37.6	21990	24	ABK50071	Human Cyp 4A11 gen
12	966	37.5	4123	24	ABK50067	DNA encoding mouse
13	534	20.7	2343	22	AA521297	Human cDNA sequenc
14	534	20.7	2343	25	ACA03656	CDNA encoding huma
15	534	20.7	2343	25	ACA04077	Human cDNA encodin
16	534	20.7	2343	25	ABK89194	DNA encoding novel
17	534	20.7	2381	24	ABK63921	CDNA encoding huma
18	534	20.7	2416	22	ABD09941	Human drug metabo
19	528.8	20.5	1894	24	ABD07693	Breast cancer-asso
20	528.8	20.5	2015	24	AA24015	Human drug metabo
21	528.8	20.5	2020	24	ABK33550	CDNA encoding huma
22	519.4	20.2	2337	24	ABK32898	CDNA encoding huma
23	512.4	19.9	2356	22	AA19563	Human expressed po
24	512.4	19.9	2356	22	ABA06504	Human cDNA SEQ ID
25	512.4	19.9	2356	22	AA540801	CDNA encoding nove
26	512.4	19.9	2356	22	AA528959	CDNA encoding for
27	512.4	19.9	2356	22	AA529592	Human endocrine po
28	512.4	19.9	2356	22	AA534859	CDNA encoding nove
29	512.4	19.9	2356	24	ABV83841	Human polynucleoti
30	512.4	19.9	2356	24	ABV07794	Novel human nuclei
31	512.4	19.9	1718	24	AB211538	Human polynucleoti
32	506.8	19.7	2084	22	AB21227	Oesophagus cancer
33	408.4	15.9	1586	22	ABD09946	Human drug metabo
34	392.2	15.2	1657	25	ACC50320	Breast cancer asso
35	387.4	15.0	17717	22	AAK82008	Human immune/haema
36	387.4	15.0	17717	22	AAK82008	Human immune/haema
37	387.4	15.0	17719	22	AAK82006	Human immune/haema
38	359.8	14.0	1508	24	AB211227	Human polynucleoti
39	320	12.4	433	25	ABK45374	Bovine BST associa
40	318.4	12.4	2073	24	ABK83594	Human cDNA diffe
41	318.4	12.4	2073	24	ABN95103	Gene #1601 used to
42	318.4	12.4	2073	24	ABL68352	Kidney cancer rela
43	318.4	12.4	2073	24	ABL68820	Kidney cancer rela
44	315.2	12.2	2211	23	AA567737	DNA encoding novel
45	315.2	12.2	2211	23	AA593586	DNA encoding novel

ALIGNMENTS

RESULT 1	ABN95640	standard; DNA; 2576 BP.
ID	ABN95640;	
AC	ABN95640;	
XX		
DT	13-AUG-2002 (first entry)	
XX		
DE	Gene #2138 used to diagnose liver cancer.	
XX		
KW	Gene, liver cancer; de; hepatocellular carcinoma; hepatotropic;	
KW	metastatic liver tumor; cytostatic; expression profile; disease state;	
KW	metastatic progression; drug toxicity; drug efficacy; drug metabolism.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200229103-A2.	
XX		
PD	11-APR-2002.	
XX		
PF	02-OCT-2001; 2001WO-US30589.	
XX		
PR	02-OCT-2000; 2000US-237054P.	
XX		
PA	(GENE-) GENE LOGIC INC.	
XX		
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;	
XX		
DR	WPI, 2002-426119/45.	
XX		
PT	Diagnosing and detecting the progression of liver cancer.	

PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -

Claim 1; SEQ ID NO 2138; 298bp; English.

CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatocytic, and
CC cytostratic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 2576 BP; 572 A; 793 C; 583 G; 628 T; 0 other;

Query Match 100.0%; Score 2576; DB 24; Length 2576;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGAGAGATCCAGAGGCTGACACATGAGTGTCTGTGAGGCCCCGACA 60
DB 1 GAATTCGGAGAGATCCAGAGGCTGACACATGAGTGTCTGTGAGGCCCCGACA 60
QY 61 GACTCCTGGAGTGTCTGTGATCTCTCAAGGCGCTCCCTGCTCAATTCGTCTGC 120
DB 61 GACTCCTGGAGTGTCTGTGATCTCTCAAGGCGCTCCCTGCTCAATTCGTCTGC 120
QY 121 TGTCTGATCAAGGCGATTCAGCTCTACCTGACAGGCGATGCTCTCAAGCCTTCAAGC 180
DB 121 TGTCTGATCAAGGCGATTCAGCTCTACCTGACAGGCGATGCTCTCAAGCCTTCAAGC 180
QY 181 AGTTCCTGGAGTGTCTGTGATCTCTCAAGGCGCTCCCTGCTCAATTCGTCTGC 240
DB 181 AGTTCCTGGAGTGTCTGTGATCTCTCAAGGCGCTCCCTGCTCAATTCGTCTGC 240
QY 241 AGGAGCTACACAGGATTCAGAAATGAGTGAAGATTCCTCAAGTCTCTCAATTCG 300
DB 241 AGGAGCTACACAGGATTCAGAAATGAGTGAAGATTCCTCAAGTCTCTCAATTCG 300
QY 301 TATGGGAGGCAAAATTCGTGTCAGCTCTATGACCTGACTATATGAGGTGATTCG 360
DB 301 TATGGGAGGCAAAATTCGTGTCAGCTCTATGACCTGACTATATGAGGTGATTCG 360
QY 361 GAGATGACAGCCGAAATTCATGAGTTCCTCAAGTCTCTCAAGTCTCTCAAGTTCG 420
DB 361 GAGATGACAGCCGAAATTCATGAGTTCCTCAAGTCTCTCAAGTCTCTCAAGTTCG 420
QY 421 GCTTGTCTCTGTTGATGAGGAGACATGATTCAGACATGACGAGATGCTCAAGCCT 480
DB 421 GCTTGTCTCTGTTGATGAGGAGACATGATTCAGACATGACGAGATGCTCAAGCCT 480
QY 481 TCCACTATGACATCTGAAAGCTCTATGAGGCTCTATGAGCTCTGATGAGTATGC 540
DB 481 TCCACTATGACATCTGAAAGCTCTATGAGGCTCTATGAGCTCTGATGAGTATGC 540
QY 541 TGAAGAAATGAGAGAGCTCTTGGCGAGATTCCTCTGAGAGTCTTTTCAAGCCT 600
DB 541 TGAAGAAATGAGAGAGCTCTTGGCGAGATTCCTCTGAGAGTCTTTTCAAGCCT 600
QY 601 CTTGATGACCTTGAAGACATCAATGAGTGTGCTTCAAGCCTTCAAGCCTTCAAGC 660
DB 601 CTTGATGACCTTGAAGACATCAATGAGTGTGCTTCAAGCCTTCAAGCCTTCAAGC 660
QY 661 TGAAGAAATTCATGCTCTATGACGAGCCTTATGAGCTTGAAGAACTGCTTTT 720

DB 661 TGAAGAAATTCATGCTCTATGACGAGCCTTATGAGCTTGAAGAACTGCTTTT 720
QY 721 CCCGTGAGGAATGCTTTTCAAGAGTACACCATCTAAGCTGACTCTGTGCGC 780
DB 721 CCCGTGAGGAATGCTTTTCAAGAGTACACCATCTAAGCTGACTCTGTGCGC 780
QY 781 GCTGACACACCGGCTGCACTGCGCCATAGGACACAGACCAAGTATCCATGTA 840
DB 781 GCTGACACACCGGCTGCACTGCGCCATAGGACACAGACCAAGTATCCATGTA 840
QY 841 GAAAGCTCAATCAAGAGAGGAGGAGCTGAGAGATCAAGAGAGAGGAGCTTTG 900
DB 841 GAAAGCTCAATCAAGAGAGGAGGAGCTGAGAGATCAAGAGAGAGGAGCTTTG 900
QY 901 ATTTTGTGATATCTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 901 ATTTTGTGATATCTCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
QY 961 ACCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
DB 961 ACCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
QY 1021 TCTCTGATCTCTATGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 1021 TCTCTGATCTCTATGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
QY 1081 AGATTCACACGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1081 AGATTCACACGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 1141 CTTACACACATGATGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
DB 1141 CTTACACACATGATGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
QY 1201 GGAAGAGCTCAGACATCCGAGTCACTCTCTGATGAGGAGGAGGAGGAGGAG 1260
DB 1201 GGAAGAGCTCAGACATCCGAGTCACTCTCTGATGAGGAGGAGGAGGAGGAG 1260
QY 1261 TGTCT 1320
DB 1261 TGTCT 1320
QY 1321 TGTCT 1380
DB 1321 TGTCT 1380
QY 1381 TCTCAGAGGATCAAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
DB 1381 TCTCAGAGGATCAAGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
QY 1441 CCAAGGCTCAGACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
DB 1441 CCAAGGCTCAGACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
QY 1501 CCAATGACGATCTGTTGATTCGAAATTCGAAATTCGAAATTCGAAATTCG 1560
DB 1501 CCAATGACGATCTGTTGATTCGAAATTCGAAATTCGAAATTCGAAATTCG 1560
QY 1561 CTAACCTCTGAG 1620
DB 1561 CTAACCTCTGAG 1620
QY 1621 GACCGGCT 1680
DB 1621 GACCGGCT 1680
QY 1681 CTTCTTCT 1740
DB 1681 CTTCTTCT 1740
QY 1741 CCAAGGCTCCTCAAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1800

Db 1741 CCAAGGCTCCCTACCTGCTTGTCTACCTGCTCTCCACCACTGATCTCTGTTGGAGAA 1800
 Qy 1801 AAAGCTGAGTGTGGAGAGAGCTGAGGCGAGCTTGACATGTGACATTAATGTAAGAAG 1860
 Db 1801 AAAGCTGAGTGTGGAGAGAGCTGAGGCGAGCTTGACATGTGACATTAATGTAAGAAG 1860
 Qy 1861 TCTTGAATATATCCAGAGATCCAGGGGCTTAAACCCCTTGTGGCCTTTGGAACAACAAC 1920
 Db 1861 TCTTGAATATATCCAGAGATCCAGGGGCTTAAACCCCTTGTGGCCTTTGGAACAACAAC 1920
 Qy 1921 TCTGTCGTAAGAGGAGGAGGCTAACCTGACGACCATTAATCAAGCCCGGGCATTAATA 1980
 Db 1921 TCTGTCGTAAGAGGAGGAGGCTAACCTGACGACCATTAATCAAGCCCGGGCATTAATA 1980
 Qy 1981 CCCCTGTCGTCGTAAGATTCAGAGGCTGTCGTCGTAAGATTCGTCGTAAGATTCGTCG 2040
 Db 1981 CCCCTGTCGTCGTAAGATTCAGAGGCTGTCGTCGTAAGATTCGTCGTAAGATTCGTCG 2040
 Qy 2041 GCTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2100
 Db 2041 GCTCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2100
 Qy 2101 CATTAATCTCAAGTAACAGAGAGATGCTTAAACCGTCAAGCTGTAATGTCGTCGTCGTA 2160
 Db 2101 CATTAATCTCAAGTAACAGAGAGATGCTTAAACCGTCAAGCTGTAATGTCGTCGTCGTA 2160
 Qy 2161 GCACATGTCCTTTCGACACCCATTCCTCAACACCTGTCGTCGTCGTCGTCGTCGTCGTCG 2220
 Db 2161 GCACATGTCCTTTCGACACCCATTCCTCAACACCTGTCGTCGTCGTCGTCGTCGTCGTCG 2220
 Qy 2221 ATAAATATCTGACCTTCAGAGCTGCGGAGCTTCACAGCTTCATCTTAAAGCTTGGCGG 2280
 Db 2221 ATAAATATCTGACCTTCAGAGCTGCGGAGCTTCACAGCTTCATCTTAAAGCTTGGCGG 2280
 Qy 2281 CCCTGAGACCTTCTCTCTCAAACTGTCCTTTCACAGCTTTCGTCGTCGTCGTCGTCGTCG 2340
 Db 2281 CCCTGAGACCTTCTCTCTCAAACTGTCCTTTCACAGCTTTCGTCGTCGTCGTCGTCGTCG 2340
 Qy 2341 TGTCAACCCCAAGACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2400
 Db 2341 TGTCAACCCCAAGACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2400
 Qy 2401 CCTCCCAAACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2460
 Db 2401 CCTCCCAAACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2460
 Qy 2461 GCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2520
 Db 2461 GCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2520
 Qy 2521 TCAATGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2576
 Db 2521 TCAATGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2576

RESULT 2

ABK50069
ID ABK50069 standard; DNA; 2576 BP.

ABK50069;

15-JUL-2002 (first entry)

XX DNA encoding human Cyp 4A11 protein.

XX Human; Cyp 4A14; hypertension; Cyp 4A11; testosterone inhibitor; obesity;

XX lipid metabolism disease; pancreatic dysfunction; type II diabetes;

XX cardiovascular disease; gene; de.

XX Chimeric - Homo sapiens.

XX Chimeric - Synthetic.

XX Key Location/Qualifiers

FT CDS 33..1592
 FT /tag= a
 FT /product= "Human Cyp 4A11 protein"
 PN WO200217856-A2.
 XX
 PD 07-MAR-2002.
 PF 29-AUG-2001; 2001WO-US26914.
 PF 29-AUG-2000; 2000US-228947P.
 PR (UYVA-) UNIV VANDERBILT.
 PA Capdevila J, Waterman M, Holia V;
 PI MPI; 2002-382929/41.
 DR P-PSDB; AAT79994.
 XX
 XX Treating hypertension in an individual by inhibiting testosterone
 PT activity, enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11
 PT activity in the individual -
 PS Claim 11; Page 71-72; 80pp; English.
 XX
 XX The present invention relates to a new method of treating hypertension in
 CC an individual. The method comprises inhibiting testosterone activity,
 CC enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the
 CC individual. Enhancing Cyp 4A14 activity leads to inhibition of
 CC testosterone activity which in turn leads to inhibition of Cyp 4A11
 CC activity. The method of the invention can be used for treating
 CC hypertension in an individual. The molecules of the invention are also
 CC useful for treating disease states associated with lipid metabolism,
 CC pancreatic dysfunction, obesity, type II diabetes, and other
 CC cardiovascular diseases. The present nucleic acid sequence encodes the
 CC human Cyp 4A11 protein of the invention, as described above.
 XX
 SQ Sequence 2576 BP; 572 A; 793 C; 583 G; 628 T; 0 other;
 Query Match 100.0%; Score 2576; DB 24; Length 2576;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GAATTCGAGAGATCCAGAGGTCGTCGACATGAGTGTCTGTGTCGAGCCCGACGA 60
 Db 1 GAATTCGAGAGATCCAGAGGTCGTCGACATGAGTGTCTGTGTCGAGCCCGACGA 60
 Qy 61 GACTCTGGGTGATGTCCTGGAATCTTCAGAGCGGCTCCGTCATTCCTGCTTGC 120
 Db 61 GACTCTGGGTGATGTCCTGGAATCTTCAGAGCGGCTCCGTCATTCCTGCTTGC 120
 Qy 121 TGTGTATCAAGGAGTCACTTACCTGACAGAGGAGTGGCTCAAGCCCTCCAGC 180
 Db 121 TGTGTATCAAGGAGTCACTTACCTGACAGAGGAGTGGCTCAAGCCCTCCAGC 180
 Qy 181 AGTTCCTGTCCTCCCTCCCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 240
 Db 181 AGTTCCTGTCCTCCCTCCCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 240
 Qy 241 AGGAGCTACACGAGATTCAGAAATGGGTGAGACATTCCTCCAGTGTGCTTCATTGGC 300
 Db 241 AGGAGCTACACGAGATTCAGAAATGGGTGAGACATTCCTCCAGTGTGCTTCATTGGC 300
 Qy 301 TATGGGAGGCAAGTGTGTCGACCTGATGACCTGATCTATGAGAGGTGATTCTGG 360
 Db 301 TATGGGAGGCAAGTGTGTCGACCTGATGACCTGATCTATGAGAGGTGATTCTGG 360
 Qy 361 GGAGATCAAGCCGGAATCCATGTCCTACAGATTCCTGTCGTCGTCGTCGTCGTCGTCGTCG 420
 Db 361 GGAGATCAAGCCGGAATCCATGTCCTACAGATTCCTGTCGTCGTCGTCGTCGTCGTCGTCG 420
 Qy 421 GCTTGTCTCTGTTGAATGGGAGACATGCTTCAGAGATGACGAGATGACCCAGCCT 480

421 GCTTGTCTCTGTGAATGGGAGACATGTTCCAGATCGAGGAGTCCAGCCAGCTT 480
481 TCACATATGACATCTGAGAGCCCTATGTGGGGCTCATATGAGAGACTCTGTACAGATATGC 540
481 TCACATATGACATCTGAGAGCCCTATGTGGGGCTCATATGAGAGACTCTGTACAGATATGC 540
541 TGGACAAATGGGAGAGCTCTTGGCCAGGATTCCTCTGTAGAGTCTTTCAGACAGTCT 600
541 TGGACAAATGGGAGAGCTCTTGGCCAGGATTCCTCTGTAGAGTCTTTCAGACAGTCT 600
541 TGGACAAATGGGAGAGCTCTTGGCCAGGATTCCTCTGTAGAGTCTTTCAGACAGTCT 600
601 CCTTGAATGACCTTGGACACCATATGAGATGCTTCAAGCATCAAGGAGAGATCCAGG 660
601 CCTTGAATGACCTTGGACACCATATGAGATGCTTCAAGCATCAAGGAGAGATCCAGG 660
601 CCTTGAATGACCTTGGACACCATATGAGATGCTTCAAGCATCAAGGAGAGATCCAGG 660
661 TGGACAGAAATCTCACTGCTTACATACAGAGCCATTAAGTGAATCAACCTGGTTTTT 720
661 TGGACAGAAATCTCACTGCTTACATACAGAGCCATTAAGTGAATCAACCTGGTTTTT 720
661 TGGACAGAAATCTCACTGCTTACATACAGAGCCATTAAGTGAATCAACCTGGTTTTT 720
721 CCGGTGTGAGGAAATGCTTTCACCAAGATGACACATCTACAGCTGACCTTGTGGCC 780
721 CCGGTGTGAGGAAATGCTTTCACCAAGATGACACATCTACAGCTGACCTTGTGGCC 780
721 CCGGTGTGAGGAAATGCTTTCACCAAGATGACACATCTACAGCTGACCTTGTGGCC 780
781 GCTGAGACACACCGGCTGCTGACCTGAGCCATCAAGGACACAGACCAAGATCCAACTGA 840
781 GCTGAGACACACCGGCTGCTGACCTGAGCCATCAAGGACACAGACCAAGATCCAACTGA 840
781 GCTGAGACACACCGGCTGCTGACCTGAGCCATCAAGGACACAGACCAAGATCCAACTGA 840
841 GGAAGGCTCAACTACAGAGAGAGGAGGCTGAGAGATCAAGAGAGAGAGGCAATTTGG 900
841 GGAAGGCTCAACTACAGAGAGAGGAGGCTGAGAGATCAAGAGAGAGAGGCAATTTGG 900
841 GGAAGGCTCAACTACAGAGAGAGGAGGCTGAGAGATCAAGAGAGAGAGGCAATTTGG 900
901 AATTTCTGATATCTCTCTCTTGGCCAAAATGAGAAATGGAGATCTTGTCAACAAGG 960
901 AATTTCTGATATCTCTCTCTTGGCCAAAATGAGAAATGGAGATCTTGTCAACAAGG 960
901 AATTTCTGATATCTCTCTCTTGGCCAAAATGAGAAATGGAGATCTTGTCAACAAGG 960
961 ACCCTGCTGAGAGTGAAGTGAACAGTTCAATTTTGAAGGCTCAACACAGCCAGTGGGA 1020
961 ACCCTGCTGAGAGTGAAGTGAACAGTTCAATTTTGAAGGCTCAACACAGCCAGTGGGA 1020
961 ACCCTGCTGAGAGTGAAGTGAACAGTTCAATTTTGAAGGCTCAACACAGCCAGTGGGA 1020
1021 TCTCCCTGAGATCTCTATGCTCTGAGGACACACCCCAAGATCAGAGAGAGGCTGGAGG 1080
1021 TCTCCCTGAGATCTCTATGCTCTGAGGACACACCCCAAGATCAGAGAGAGGCTGGAGG 1080
1021 TCTCCCTGAGATCTCTATGCTCTGAGGACACACCCCAAGATCAGAGAGAGGCTGGAGG 1080
1081 AGATTCACAGCTCTCTGAGTGAATGAGAGCTTCATCACTTGGAAACCACTTGGACCAATGC 1140
1081 AGATTCACAGCTCTCTGAGTGAATGAGAGCTTCATCACTTGGAAACCACTTGGACCAATGC 1140
1081 AGATTCACAGCTCTCTGAGTGAATGAGAGCTTCATCACTTGGAAACCACTTGGACCAATGC 1140
1141 CTTACACACATGATGATTAAGAGGACCTGAGGCTTCACACCGGTGCTCAGGCAATTC 1200
1141 CTTACACACATGATGATTAAGAGGACCTGAGGCTTCACACCGGTGCTCAGGCAATTC 1200
1141 CTTACACACATGATGATTAAGAGGACCTGAGGCTTCACACCGGTGCTCAGGCAATTC 1200
1201 GCGAGAGCTCAAGATCTCCGCTCACTTCCCTGATGAGGCGCTCTTGGCCAAAGATCA 1260
1201 GCGAGAGCTCAAGATCTCCGCTCACTTCCCTGATGAGGCGCTCTTGGCCAAAGATCA 1260
1201 GCGAGAGCTCAAGATCTCCGCTCACTTCCCTGATGAGGCGCTCTTGGCCAAAGATCA 1260
1261 TGGTCTCTCTCTCAATTAATGAGCTTCAACCAACCAAGATGAGGCTTCCAGAGG 1320
1261 TGGTCTCTCTCTCAATTAATGAGCTTCAACCAACCAAGATGAGGCTTCCAGAGG 1320
1261 TGGTCTCTCTCTCAATTAATGAGCTTCAACCAACCAAGATGAGGCTTCCAGAGG 1320
1321 TGTATGACCTTTCCTGTTTGGACCGGCTTCTGCTCAACACAGGCAAGCTTTCCTGCTT 1380
1321 TGTATGACCTTTCCTGTTTGGACCGGCTTCTGCTCAACACAGGCAAGCTTTCCTGCTT 1380
1321 TGTATGACCTTTCCTGTTTGGACCGGCTTCTGCTCAACACAGGCAAGCTTTCCTGCTT 1380
1321 TGTATGACCTTTCCTGTTTGGACCGGCTTCTGCTCAACACAGGCAAGCTTTCCTGCTT 1380
1381 TCTGAGAGAGATCAAGAGATCTGATTTGGAGAAACAATTTGCAATGAAAGAGTGAAGTGG 1440
1381 TCTGAGAGAGATCAAGAGATCTGATTTGGAGAAACAATTTGCAATGAAAGAGTGAAGTGG 1440
1381 TCTGAGAGAGATCAAGAGATCTGATTTGGAGAAACAATTTGCAATGAAAGAGTGAAGTGG 1440
1441 CCAAGGCTGAGACCTGCTCCGCTTGGAGTGTGCTGATCCCAAGAGATCCCAATCC 1500
1441 CCAAGGCTGAGACCTGCTCCGCTTGGAGTGTGCTGATCCCAAGAGATCCCAATCC 1500
1441 CCAAGGCTGAGACCTGCTCCGCTTGGAGTGTGCTGATCCCAAGAGATCCCAATCC 1500
1501 CCATTTGACAGATTTGTGTTGAATTCMAAAATGGAATCACCTGCTCAGAGAGGCTCC 1560
1501 CCATTTGACAGATTTGTGTTGAATTCMAAAATGGAATCACCTGCTCAGAGAGGCTCC 1560

1561 CTAACTCTTGTGAAGACAAAGACCAAGCTTGTAGGGCTCCACCTGCTCTGCTTCT 1620
1561 CTAACTCTTGTGAAGACAAAGACCAAGCTTGTAGGGCTCCACCTGCTCTGCTTCTCT 1620
1621 GACCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
1621 GACCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
1621 GACCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
1681 CCTTCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1681 CCTTCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1681 CCTTCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1741 CCAAGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1741 CCAAGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1741 CCAAGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
1801 AAAGCTGAGTGTGAGGAG 1860
1801 AAAGCTGAGTGTGAGGAG 1860
1801 AAAGCTGAGTGTGAGGAG 1860
1861 TCTTGAATCATGTCAGAGATCCAGGCTTAAACCCCTTGTGCTTGGAAACCAAGC 1920
1861 TCTTGAATCATGTCAGAGATCCAGGCTTAAACCCCTTGTGCTTGGAAACCAAGC 1920
1861 TCTTGAATCATGTCAGAGATCCAGGCTTAAACCCCTTGTGCTTGGAAACCAAGC 1920
1921 TCTGCTGAGAGGAGTGAAGAGCTAACCTGACGACCAATATCTTAAGCCCGGAGATAAA 1980
1921 TCTGCTGAGAGGAGTGAAGAGCTAACCTGACGACCAATATCTTAAGCCCGGAGATAAA 1980
1921 TCTGCTGAGAGGAGTGAAGAGCTAACCTGACGACCAATATCTTAAGCCCGGAGATAAA 1980
1981 CCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
1981 CCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
1981 CCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
2041 GCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2041 GCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2041 GCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2101 CATTAATCAAGTGAACAG 2160
2101 CATTAATCAAGTGAACAG 2160
2101 CATTAATCAAGTGAACAG 2160
2161 GCAACATGCTTTCAG 2220
2161 GCAACATGCTTTCAG 2220
2161 GCAACATGCTTTCAG 2220
2221 ATAAATATGAG 2280
2221 ATAAATATGAG 2280
2221 ATAAATATGAG 2280
2281 CCTGAG 2340
2281 CCTGAG 2340
2281 CCTGAG 2340
2341 TGTACACCCCAAG 2400
2341 TGTACACCCCAAG 2400
2341 TGTACACCCCAAG 2400
2441 TGTACACCCCAAG 2460
2441 TGTACACCCCAAG 2460
2441 TGTACACCCCAAG 2460
2461 GCTGAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2461 GCTGAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2461 GCTGAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2521 TCACTGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2576
2521 TCACTGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2576

RESULT 3
ABL68591
ID ABL68591 standard; DNA; 2576 BP.

XX	ABL68591;
AC	
XX	
XX	15-MAY-2002 (first entry)
DE	Kidney cancer related gene sequence SEQ ID NO:6928.
XX	
XX	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinous;
KM	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX	gene; ds.
OS	Homo sapiens.
PX	
PN	MO200194629-A2.
PD	
PF	13-DEC-2001.
PP	
PR	30-MAY-2001; 2001KO-US10838.
PT	
XX	
XX	05-JUN-2000; 2000US-209473P.
XX	05-JUN-2000; 2000US-209531P.
XX	18-SEP-2000; 2000US-231133P.
XX	18-SEP-2000; 2000US-233617P.
XX	20-SEP-2000; 2000US-234003P.
XX	20-SEP-2000; 2000US-234034P.
XX	20-SEP-2000; 2000US-234052P.
XX	22-SEP-2000; 2000US-234507P.
XX	22-SEP-2000; 2000US-234567P.
XX	25-SEP-2000; 2000US-234923P.
XX	25-SEP-2000; 2000US-234924P.
XX	25-SEP-2000; 2000US-235077P.
XX	25-SEP-2000; 2000US-235082P.
XX	25-SEP-2000; 2000US-235134P.
XX	25-SEP-2000; 2000US-235280P.
XX	26-SEP-2000; 2000US-235637P.
XX	26-SEP-2000; 2000US-235638P.
XX	27-SEP-2000; 2000US-235711P.
XX	27-SEP-2000; 2000US-235720P.
XX	27-SEP-2000; 2000US-235840P.
XX	27-SEP-2000; 2000US-235863P.
XX	28-SEP-2000; 2000US-236028P.
XX	28-SEP-2000; 2000US-236032P.
XX	28-SEP-2000; 2000US-236033P.
XX	28-SEP-2000; 2000US-236034P.
XX	28-SEP-2000; 2000US-236109P.
XX	28-SEP-2000; 2000US-236111P.
XX	29-SEP-2000; 2000US-236842P.
XX	29-SEP-2000; 2000US-236891P.
XX	02-OCT-2000; 2000US-237172P.
XX	02-OCT-2000; 2000US-237173P.
XX	02-OCT-2000; 2000US-237276P.
XX	02-OCT-2000; 2000US-237294P.
XX	02-OCT-2000; 2000US-237295P.
XX	02-OCT-2000; 2000US-237316P.
XX	03-OCT-2000; 2000US-237425P.
XX	03-OCT-2000; 2000US-237598P.
XX	03-OCT-2000; 2000US-237604P.
XX	03-OCT-2000; 2000US-237605P.
XX	03-OCT-2000; 2000US-237608P.
XX	01-NOV-2000; 2000US-244867P.
XX	01-NOV-2000; 2000US-245084P.
XX	
PA	(AVAL-) AVALON PHARM.
PI	
PI	Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S,
PI	Soppet DR, Weaver Z;
XX	
XX	WPI; 2002-188264/24.
XX	
XX	Screening for anti-neoplastic agent involves exposing cells to a
XX	chemical agent to be tested for anti-neoplastic activity, and
XX	determining a change in expression of a gene of a signature gene set

XX	PS	Claim 1; SEQ ID 6928; 44pp; English.
CC	XX	The present invention describes a method (M1) for screening for an
CC	XX	anti-neoplastic agent. The method involves exposing cells to a chemical
CC	XX	agent to be tested for anti-neoplastic activity, determining a change in
CC	XX	expression of at least one gene (I) of a signature gene set, where (I)
CC	XX	comprises a sequence (S) selected from 847 sequences (given in AB161664
CC	XX	to AB170110), of is at least 95% identical to (S), where a change in
CC	XX	expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC	XX	activity and can be used in gene therapy. M1 can be used for screening
CC	XX	an anti-neoplastic agent, and can be used for producing a product which
CC	XX	is the data collected with respect to the anti-neoplastic agent as a
CC	XX	result of M1, and the data is sufficient to convey the chemical
CC	XX	structure and/or properties of the agent. M1 can be used in the
CC	XX	treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC	XX	oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC	XX	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC	XX	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC	XX	carcinoma, papillary carcinoma and Wilms' tumour.
XX	XX	
XX	XX	Sequence 2576 BP; 572 A; 793 C; 583 G; 628 T; 0 other;
XX	XX	
XX	XX	Query Match 100.0%; Score 2576; DB 24; Length 2576;
XX	XX	Best Local Similarity 100.0%; Pred. No. 0;
XX	XX	Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	DB	1 GAATTCGCGAGAGATCCAGCAGGTGTCGACACATGATGTCTGTGCTGAGCCCCACACA 60
DB	1	GAATTCGCGAGAGATCCAGCAGGTGTCGACACATGATGTCTGTGCTGAGCCCCACACA 60
QY	61	GACTCTGGGTGATGTCTTGGAATCCTCCAGGGGCTCCCTGCTCATTTGCTTTGCG 120
DB	61	GACTCTGGGTGATGTCTTGGAATCCTCCAGGGGCTCCCTGCTCATTTGCTTTGCG 120
QY	121	TGCGATCAAGGGCAGTTCAGCTCAACGTGACAGGAGTGGCTGCTCAAGGCCCTCCAGC 180
DB	121	TGCGATCAAGGGCAGTTCAGCTCAACGTGACAGGAGTGGCTGCTCAAGGCCCTCCAGC 180
QY	181	AGTTCGCGTGCCTCCCTCCCATCTGAGCTCTTGCGGACATCCAGAGGCTCCAAACAGACC 240
DB	181	AGTTCGCGTGCCTCCCTCCCATCTGAGCTCTTGCGGACATCCAGAGGCTCCAAACAGACC 240
QY	241	AGAGCTACACCGGATTCAGAAATGGGGTGAGACATTTCCAAATGTCCTGCTCATTTGGC 300
DB	241	AGAGCTACACCGGATTCAGAAATGGGGTGAGACATTTCCAAATGTCCTGCTCATTTGGC 300
QY	301	TATGGGAGGCGAAGTTCTGTGTCAGGCTTATGACCCGTATATATGAAAGGATTCGCG 360
DB	301	TATGGGAGGCGAAGTTCTGTGTCAGGCTTATGACCCGTATATATGAAAGGATTCGCG 360
QY	361	GGAGATCAGACCCGAAATCCCATGATTCCTACATTTCTGGCTCCATGATTTGGGTAACG 420
DB	361	GGAGATCAGACCCGAAATCCCATGATTCCTACATTTCTGGCTCCATGATTTGGGTAACG 420
QY	421	GCTTGTCTCTGTGGAATGGGCGACATGGTTCCAGCATGACGGAATGCTGACCCGAGCTT 480
DB	421	GCTTGTCTCTGTGGAATGGGCGACATGGTTCCAGCATGACGGAATGCTGACCCGAGCTT 480
QY	481	TCCACTATGACATCTGGAAGCCCTATATGTTGGGGCTCATAGGCACTCTGTACGAATGAC 540
DB	481	TCCACTATGACATCTGGAAGCCCTATATGTTGGGGCTCATAGGCACTCTGTACGAATGAC 540
QY	541	TGAGCAATGGGAAGAGCTCCTTGGCCGAGATTCCTCCGTAAGGCTCTTTCAGACGCTCT 600
DB	541	TGAGCAATGGGAAGAGCTCCTTGGCCGAGATTCCTCCGTAAGGCTCTTTCAGACGCTCT 600
QY	601	CTTGAATGACCTTGACACCATCATGAAATGTGCTTCAGGCATTCAGGGCAGCATTCAGG 660
DB	601	CTTGAATGACCTTGACACCATCATGAAATGTGCTTCAGGCATTCAGGGCAGCATTCAGG 660
QY	661	TGAGCAGGAATTCAGTCTCATATACAGGCGATTAGTGAACCTGAAACAACCTGGATTCTT 720
DB	661	TGAGCAGGAATTCAGTCTCATATACAGGCGATTAGTGAACCTGAAACAACCTGGATTCTT 720

XX 02-OCT-2001; 2001WO-US30589.
 XX 02-OCT-2000; 2000US-237054P.
 XX (GENE-) GENE LOGIC INC.
 XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX WPI; 2002-426119/45.
 DR
 XX
 PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 PS
 PS Claim 1; SEQ ID NO 1610; 298bp; English.
 XX
 XX The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cyostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 1763 BP; 381 A; 538 C; 430 G; 414 T; 0 other;
 Query Match 63.1%; Score 1625.8; DB 24; Length 1763;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;

DB 481 ACATCTGAAGGCTTATGTGGGGCTATGGCAACCTCTGTACAGATGATCTGGACAAAT 540
 QY
 QY 550 GGAAGAGCTCTCTGGCCAGATTCCTCTGGAGGCTTTTACAGAGCTCTCTTGA 609
 DB 541 GGAAGAGCTCTCTGGCCAGATTCCTCTGGAGGCTTTTACAGAGCTCTCTTGA 600
 QY 610 CCTGGACACCATCATGAAGTGTCTTACGACCATGAGGAGACATTCAGGTGACAGA 669
 DB 601 CCTGGACACCATCATGAAGTGTCTTACGACCATGAGGAGACATTCAGGTGACAGA 660
 QY 670 ATTTCAGGCTTACATACAGGCGCATTAAGTGAACCTGAACACCTGGTTTTTCCGCTGA 729
 DB 661 ATTTCAGGCTTACATACAGGCGCATTAAGTGAACCTGAACACCTGGTTTTTCCGCTGA 720
 QY 730 GGAATGCTTTTACAGAGATGACACATCTACAGCTGACCTTGTGCGCGCTGACAC 789
 DB 721 GGAATGCTTTTACAGAGATGACACATCTACAGCTGACCTTGTGCGCGCTGACAC 780
 QY 790 ACCGCGCTGGCAGCTGGCCCATCAGACAC--AGACCAAGTATCCATGAGAGAG 846
 DB 781 ACCGCGCTGGCAGCTGGCCCATCAGACACCAAGATGATCCAACTGAGAGAG 840
 QY 847 CTCACTACAGAGAGGGGGAGCTGAGAGATCAAGAGAGAGAGCATTTGATTTTC 906
 DB 841 CTCACTACAGAGAGGGGGAGCTGAGAGATCAAGAGAGAGAGCATTTGATTTTC 900
 QY 907 TGAATATCTCTCTTGGCCAAATAGAGAAATGGAGACATCTGTGACACAGACCTCC 966
 DB 901 TGAATATCTCTCTTGGCCAAATAGAGAAATGGAGACATCTGTGACACAGACCTCC 960
 QY 967 GTGTGAGGTGACACAGTTCATGTTTGAAGGCCACACACACAGCCAGTGGATCTCT 1026
 DB 961 GTGTGAGGTGACACAGTTCATGTTTGAAGGCCACACACACAGCCAGTGGATCTCT 1020
 QY 1027 GATCTCTATATGCTGTGGCCACACACCCAGATCAGAGAGGTGCGGAGAGATCC 1086
 DB 1021 GATCTCTATATGCTGTGGCCACACACCCAGATCAGAGAGGTGCGGAGAGATCC 1080
 QY 1087 ACAGGCTCTCTGGGTATGAGAGCTCATCATCCTGGAACCACTGGACACAGATGCCATCA 1146
 DB 1081 ACAGGCTCTCTGGGTATGAGAGCTCATCATCCTGGAACCACTGGACACAGATGCCATCA 1140
 QY 1147 CCACATGTGATTTAAGAGGACATGAGGCTTACCAACCGGTGCCAGGATGGCAGAG 1206
 DB 1141 CCACATGTGATTTAAGAGGACATGAGGCTTACCAACCGGTGCCAGGATGGCAGAG 1200
 QY 1207 AGCTGACATCCCGTCACTTCTCTGATGAGGCGCTCTTGGCCAAAGTATCATGGTCC 1286
 DB 1201 AGCTGACATCCCGTCACTTCTCTGATGAGGCGCTCTTGGCCAAAGTATCATGGTCC 1280
 QY 1267 TCTCTCTATTAATGAGGCTTACCAACCAAGTGGGCCCAACCGAGAGGTGTTG 1326
 DB 1261 TCTCTCTATTAATGAGGCTTACCAACCAAGTGGGCCCAACCGAGAGGTGTTG 1320
 QY 1327 ACCCTTCCGTTTGGACCGGGTCTGCTCAACACAGCCAGCTTCTCTGAG 1386
 DB 1321 ACCCTTCCGTTTGGACCGGGTCTGCTCAACACAGCCAGCTTCTCTGAG 1380
 QY 1387 GAGATGAAGAACTGATTTGGAAACAATTTGCCATGAAGAGTGAAGTGGCC--A 1443
 DB 1381 GAGATGAAGAACTGATTTGGAAACAATTTGCCATGAAGAGTGAAGTGGCCAGC 1440
 QY 1444 CGGCGCTGACCGGTGCGCTTGAAGCTGTGATCCACACAGATCCCATCCCA 1503
 DB 1441 AGGCGCTGACCGGTGCGCTTGAAGCTGTGATCCACACAGATCCCATCCCA 1500
 QY 1504 TTGCAACACTTGTGTTGAATCCAAATAGAAATCACTGCTCAGAGAGCTCCCTA 1563
 DB 1501 TTGCAACACTTGTGTTGAATCCAAATAGAAATCACTGCTCAGAGAGCTCCCTA 1560
 QY 1564 ACCCTTGAAGACAGAGCAAGCTTTGAGGCGCTCACTGCTGCTCTTCTGAC 1623

Db	1561	ACCCTGTGAAGACAAAGACAGCTTTTGAAGGGCCCTCCACCTGCGCTCTGTCCTCTGAC	1620
Qy	1524	CCCCGCTTCGTGCCCCCTCTGTCGACCATATCCGTCTTCTGCTGSCCACCCTCCCT	1683
Db	1621	CCCCGCTTCGTGCCCCCTCTGTCGACCATATCCGTCTTCTGCTGSCCATTACCTTG	1680
Qy	1684	TCCTTC	1688
Db	1681	TCCTCC	1685

RESULT 5
ABL68271
ID ABL68271 standard; DNA; 1763 BP

PR	05-JUN-2000	2000US-209413P.
PR	05-JUN-2000	2000US-209531P.
PR	18-SEP-2000	2000US-233133P.
PR	18-SEP-2000	2000US-233617P.
PR	20-SEP-2000	2000US-234009P.
PR	20-SEP-2000	2000US-234034P.
PR	20-SEP-2000	2000US-234052P.
PR	22-SEP-2000	2000US-234509P.
PR	22-SEP-2000	2000US-234567P.
PR	25-SEP-2000	2000US-234923P.
PR	25-SEP-2000	2000US-234924P.
PR	25-SEP-2000	2000US-235077P.
PR	25-SEP-2000	2000US-235082P.
PR	25-SEP-2000	2000US-235134P.
PR	25-SEP-2000	2000US-235280P.
PR	26-SEP-2000	2000US-235637P.
PR	26-SEP-2000	2000US-235638P.
PR	27-SEP-2000	2000US-235711P.
PR	27-SEP-2000	2000US-235720P.
PR	27-SEP-2000	2000US-235840P.
PR	27-SEP-2000	2000US-235863P.
PR	28-SEP-2000	2000US-236028P.
PR	28-SEP-2000	2000US-236032P.
PR	28-SEP-2000	2000US-236033P.
PR	28-SEP-2000	2000US-236034P.
PR	28-SEP-2000	2000US-236109P.
PR	28-SEP-2000	2000US-236111P.
PR	29-SEP-2000	2000US-236842P.
PR	29-SEP-2000	2000US-236891P.
PR	02-OCT-2000	2000US-237172P.
PR	02-OCT-2000	2000US-237173P.
PR	02-OCT-2000	2000US-237278P.
PR	02-OCT-2000	2000US-237294P.
PR	02-OCT-2000	2000US-237295P.
PR	02-OCT-2000	2000US-237316P.
PR	03-OCT-2000	2000US-237425P.
PR	03-OCT-2000	2000US-237598P.
PR	03-OCT-2000	2000US-237604P.
PR	03-OCT-2000	2000US-237606P.

03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppe DR, Weaver Z;
XX WPI; 2002-186264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1, SEQ ID 6606; 44pp: English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABLE1664 to AB170110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytosolic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 1763 BP; 381 A; 538 C; 430 G; 414 T; 0 other;

Query Match	63.14	Score 1625.8	DB 24	Length 1763
Best Local Similarity	98.64	Pred. No. 0		
Matches 1662; Conservative	0	Mismatches 17	Indels 6	Gaps 2

QY	10	AGAGATTCAGAGGAGGTCTGCACCATGATGATCTCTGTCTGAGACCCCAAGCAATCTCTGG	69
Db	1	AGAGATTCAGAGGAGGTCTGCACCATGATGATCTCTGTCTGAGACCCCAAGCAATCTCTGG	60
QY	70	GTCATGTCTCTGGAATTCCTCCAGGGGCTCCCTGCTCATATTCGTCTCTGCTGCTATCA	129
Db	61	GTCATGTCTCTGGAATTCCTCCAGGGGCTCCCTGCTCATATTCGTCTCTGCTGCTATCA	120
QY	130	AGGCAATTCAGCTCTACCTCTGCACAGGCAATGGCTGCTCAAACCCCTCCAGCAAGTCCCGT	189
Db	121	AGGCAATTCAGCTCTACCTCTGCACAGGCAATGGCTGCTCAAACCCCTCCAGCAAGTCCCGT	180
QY	190	GCCCTCCCTCCCACTGAGCTCTTTCGGGACATATCCAGAGCTCCCAACAGAACCAAGACTAC	249
Db	181	GCCCTCCCTCCCACTGAGCTCTTTCGGGACATATCCAGAGCTCCCAACAGAACCAAGACTAC	240
QY	250	AACGGATTCAGAAATGGGTGGAGACATTTCCAAATGCTGTCTCTATTGGCTATTGGGGAG	309
Db	241	AACGGATTCAGAAATGGGTGGAGACATTTCCAAATGCTGTCTCTATTGGCTATTGGGGAG	300
QY	310	GCAAAATTCGATGTCAGGCTCTATGACCCCTGACCTATATGAAAGGTGATTTCTGGGAGATCAG	369
Db	301	GCAAAATTCGATGTCAGGCTCTATGACCCCTGACCTATATGAAAGGTGATTTCTGGGAGATCAG	360
QY	370	AACCCGAATTCCTCATGATTTCTGAGCTTCATGATGATTTGGGTACGGCTTGCTCC	429
Db	361	AACCCGAATTCCTCATGATTTCTGAGCTTCATGATGATTTGGGTACGGCTTGCTCC	420
QY	430	TGTTGATATGGGAGCATGATTTCCAGCATCCAGCGAATGCTGACCCCAAGCTTCCATCATG	489
Db	421	TGTTGATATGGGAGCATGATTTCCAGCATCCAGCGAATGCTGACCCCAAGCTTCCATCATG	480

490 ACATCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACAGATGATCTGGAAGAT 549
481 ACATCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACAGATGATCTGGAAGAT 540
550 GGGAAAGAGCTCTTGGCCAGAGATTCCTCTGAGAGTCTTTTGACAGAGTCTCTTGATGA 609
541 GGGAAAGAGCTCTTGGCCAGAGATTCCTCTGAGAGTCTTTTGACAGAGTCTCTTGATGA 600
610 CCTTGAACACCATCATGAAGTGTGCTTCAAGCATCAAGGACAGATTCAGAGTGAAGGA 669
601 CCTTGAACACCATCATGAAGTGTGCTTCAAGCATCAAGGACAGATTCAGAGTGAAGGA 660
670 ATTCTCAGTCTCATATACAGAGCCATTAGTGAAGCTTGAACCAAGCTGTTTTTCCCGTGA 729
661 ATTCTCAGTCTCATATACAGAGCCATTAGTGAAGCTTGAACCAAGCTGTTTTTCCCGTGA 720
730 GGAATGCTCTTTCACAGATGACACCATCTACAGCTGACCTTCTGTCGCGCTGACAC 789
721 GGAATGCTCTTTCACAGATGACACCATCTACAGCTGACCTTCTGTCGCGCTGACAC 780
790 ACCGCGCTGCGACGCTGCGCCATCAGCAGAC--AGACCAAGTATCAACTGAGGAAG 846
781 ACCGCGCTGCGACGCTGCGCCATCAGCAGACCCAAAGTATCAACTGAGGAAG 840
847 CTCAACTACAGAAAGAGGGGAGCTGAGAAATCAAGAGGAAGGCAATTTGATTTTC 906
841 CTCAACTACAGAAAGAGGGGAGCTGAGAAATCAAGAGGAAGGCAATTTGATTTTC 900
907 TGGATATCTCTCTTGGCCAAATAGAGATGGAGCATCTTGTACAGACAGACCTCC 966
901 TGGATATCTCTCTTGGCCAAATAGAGATGGAGCATCTTGTACAGACAGACCTCC 960
967 GTGCTGAGGTGACACGTTTATGTTTGAAGGCGCCAGACACACAGCCAGTGGATCTCT 1026
961 GTGCTGAGGTGACACGTTTATGTTTGAAGGCGCCAGACACACAGCCAGTGGATCTCT 1020
1027 GGAATCTCTATGCTCTGCGCCACACCCAGATCAAGAGAGTTCGCGAGAGATTC 1086
1021 GGAATCTCTATGCTCTGCGCCACACCCAGATCAAGAGAGTTCGCGAGAGATTC 1080
1087 ACAGGCTCTGAGGTGAGAGGCTCCATCAGCTGAGACACCTGAGACAGATTCCTTCA 1146
1081 ACAGGCTCTGAGGTGAGAGGCTCCATCAGCTGAGACACCTGAGACAGATTCCTTCA 1140
1147 CCACCATGTGATTAAGAGGCACTGAGGCTCTACCAAGGTCAGAGCATTTGGCAGAG 1206
1141 CCACCATGTGATTAAGAGGCACTGAGGCTCTACCAAGGTCAGAGCATTTGGCAGAG 1200
1207 AGCTCAGCACTCCCGTCACTTCCCTGATGAGGCGCTCTTGGCCAAAGTATCATGCTCC 1266
1201 AGCTCAGCACTCCCGTCACTTCCCTGATGAGGCGCTCTTGGCCAAAGTATCATGCTCC 1260
1267 TCCTCTCAATTAAGGCGCTTCAACACCAACCAAAATGAGGCGCCCAAGAGGTCTTG 1326
1261 TCCTCTCAATTAAGGCGCTTCAACACCAACCAAAATGAGGCGCCCAAGAGGTCTTG 1320
1327 ACCCTTTCGTTTGGACCGGGTCTGCTCAACAGCAGCGTCTTCTGCTCTCAG 1386
1321 ACCCTTTCGTTTGGACCGGGTCTGCTCAACAGCAGCGTCTTCTGCTCTCAG 1380
1387 GAGATCAAGAACTGATTTGGGAAACAAATTTGCCATGAACGAGCTGAAGGTGGCAGC 1443
1381 GAGATCAAGAACTGATTTGGGAAACAAATTTGCCATGAACGAGCTGAAGGTGGCAGC 1440
1444 CGGCGCTGACCTGCTCGCGCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1503
1441 AGGCGCTGACCTGCTCGCGCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
1504 TTGACAGACTTGTGTTGAATCCAAATAGAAATCAACCTGAGCTCTGAGAGGCTCTCTA 1563
1501 TTGACAGACTTGTGTTGAATCCAAATAGAAATCAACCTGAGCTCTGAGAGGCTCTCTA 1560
1564 ACCCTTGAAGACAAAGACAGCTTGAAGGCGCTCAGCTGCGTCTGTTCTTCTGAC 1623

1561 ACCCTTGAAGACAAAGACAGCTTGAAGGCGCTCAGCTGCGTCTTCTGAC 1620
1624 CCGCGCTTGTGCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683
1621 CCGCGCTTGTGCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
1684 TCTTC 1688
1681 TCTTC 1685
RESULT 6
ABL68599 standard; DNA; 1763 BP.
ABL68599;
15-MAY-2002 (first entry)
Kidney cancer related gene sequence SEQ ID NO: 6936.
Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
gene; de.
Homo sapiens.
MO200194629-A2.
13-DEC-2001.
30-MAY-2001; 2001WO-US10838.
05-JUN-2000; 2000US-209473P.
18-SEP-2000; 2000US-209531P.
18-SEP-2000; 2000US-233133P.
20-SEP-2000; 2000US-233617P.
20-SEP-2000; 2000US-234009P.
20-SEP-2000; 2000US-234034P.
20-SEP-2000; 2000US-234052P.
22-SEP-2000; 2000US-234509P.
22-SEP-2000; 2000US-234567P.
25-SEP-2000; 2000US-234923P.
25-SEP-2000; 2000US-234924P.
25-SEP-2000; 2000US-235077P.
25-SEP-2000; 2000US-235082P.
25-SEP-2000; 2000US-235134P.
25-SEP-2000; 2000US-235280P.
26-SEP-2000; 2000US-235637P.
26-SEP-2000; 2000US-235638P.
27-SEP-2000; 2000US-235711P.
27-SEP-2000; 2000US-235720P.
27-SEP-2000; 2000US-235840P.
27-SEP-2000; 2000US-235863P.
28-SEP-2000; 2000US-236028P.
28-SEP-2000; 2000US-236032P.
28-SEP-2000; 2000US-236033P.
28-SEP-2000; 2000US-236034P.
28-SEP-2000; 2000US-236109P.
28-SEP-2000; 2000US-236111P.
29-SEP-2000; 2000US-236842P.
29-SEP-2000; 2000US-236891P.
02-OCT-2000; 2000US-237172P.
02-OCT-2000; 2000US-237173P.
02-OCT-2000; 2000US-237278P.
02-OCT-2000; 2000US-237294P.
02-OCT-2000; 2000US-237295P.
02-OCT-2000; 2000US-237316P.
03-OCT-2000; 2000US-237425P.
03-OCT-2000; 2000US-237588P.
03-OCT-2000; 2000US-237604P.

PR 03-OCT-2000: 2000US-237608P.
 PR 03-OCT-2000: 2000US-237608P.
 PR 01-NOV-2000: 2000US-244867P.
 PR 01-NOV-2000: 2000US-245084P.
 XX
 XX
 PA (AVAIL-) AVAILON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
 PI Soppet DR, Weaver Z;
 XX WPI: 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1; SEQ ID 6936; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 CC
 XX Sequence 1763 BP; 381 A; 538 C; 430 G; 414 T; 0 other;
 SQ
 Query Match 63.1%; Score 1625.8; DB 24; Length 1763;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;

QY 490 ACATCTGAAGCCCTTATGTGGGCTCATGGCAAGCTTGTACAGATATGCTGACAAAT 549
 DB 481 ACATCTGAAGCCCTTATGTGGGCTCATGGCAAGCTTGTACAGATATGCTGACAAAT 540
 QY 550 GGAAGAGCTCTTGGCAGAGATTCCTCTGAGAGTCTTTCAGACAGTCTCTTGATGA 609
 DB 541 GGAAGAGCTCTTGGCAGAGATTCCTCTGAGAGTCTTTCAGACAGTCTCTTGATGA 600
 QY 610 CCTGACACCATCATGAAGTGTCTTCAGCCATCAGGAGGAGCATCAGGTGACAGA 669
 DB 601 CCTGACACCATCATGAAGTGTCTTCAGCCATCAGGAGGAGCATCAGGTGACAGA 660
 QY 670 ATTCTAGTCTTATCATACAGGCTTATGATGCTGAAACCTGTTTTCCTGCTGA 729
 DB 661 ATTCTAGTCTTATCATACAGGCTTATGATGCTGAAACCTGTTTTCCTGCTGA 720
 QY 730 GAATGCTCTTACACGAATGACACCATCTACAGCTTACCTGCTGCGCTGACAC 789
 DB 721 GAATGCTCTTACACGAATGACACCATCTACAGCTTACCTGCTGCGCTGACAC 780
 QY 790 ACCGCGCTGCGCAGCTGCGCCATCAGACAC--AGACCAAGTATCAACTGAGAGG 846
 DB 781 ACCGCGCTGCGCAGCTGCGCCATCAGACACCAAGTATCAACTGAGAGG 840
 QY 847 CTCACTTACAGAGAGGGGAGCTGAGAAATCAAGAGAGAGGCAATTTGATTTTC 906
 DB 841 CTCACTTACAGAGAGGGGAGCTGAGAAATCAAGAGAGAGGCAATTTGATTTTC 900
 QY 907 TGGATATCCCTCTTGGCCAAATGAGAGTGGAGCATCTTGTGACAGAGACCTTC 966
 DB 901 TGGATATCCCTCTTGGCCAAATGAGAGTGGAGCATCTTGTGACAGAGACCTTC 960
 QY 967 GTGCTGAGGTGACACAGTTTATGTTGAGGAGCAACACACAGAGAGTGGATCTCT 1026
 DB 961 GTGCTGAGGTGACACAGTTTATGTTGAGGAGCAACACAGAGAGTGGATCTCT 1020
 QY 1027 GATTCCTTATGCTTGGCCACACACCCAGCATCAGAGAGTGGCGGAGAGATTC 1086
 DB 1021 GATTCCTTATGCTTGGCCACACCCAGCATCAGAGAGTGGCGGAGAGATTC 1080
 QY 1087 ACAGCCTCCGAGTATGAGAGCTTCATCACTGGAACCACTGGAACAGATGCTTACA 1146
 DB 1081 ACAGCCTCCGAGTATGAGAGCTTCATCACTGGAACCACTGGAACAGATGCTTACA 1140
 QY 1147 CCACATGTCATTAAGAGAGGCACTGAGGCTTACCAACCGGTGCCAGGCAATTTGAGAG 1206
 DB 1141 CCACATGTCATTAAGAGAGGCACTGAGGCTTACCAACCGGTGCCAGGCAATTTGAGAG 1200
 QY 1207 AGCTCAGACCTCCGATCACTTCTGTATGGGCGCTCTTGGCCAAAGTATCATGTGTC 1266
 DB 1201 AGCTCAGACCTCCGATCACTTCTGTATGGGCGCTCTTGGCCAAAGTATCATGTGTC 1260
 QY 1267 TCTCTTCATTAATGAGCTTCAACCAACCAAAAGTGTGGCCAAACCGAGAGTGTG 1326
 DB 1261 TCTCTTCATTAATGAGCTTCAACCAACCAAAAGTGTGGCCAAACCGAGAGTGTG 1320
 QY 1327 ACCCTTTCGTTTTCAGACCGGTTCTGCTCAACACAGCAGCTTCTCTGCTTTCAG 1386
 DB 1321 ACCCTTTCGTTTTCAGACCGGTTCTGCTCAACACAGCAGCTTCTCTGCTTTCAG 1380
 QY 1387 GAGATTCAGAGAACTGATTTGGAGAAACAATTTGCAATGAAGAGTGAAGTGC--A 1443
 DB 1381 GAGATTCAGAGAACTGATTTGGAGAAACAATTTGCAATGAAGAGTGAAGTGCAGC 1440
 QY 1444 CGGCGCTGACCTGCTCCGTTTGAAGTGTGCTGTATCCACAGAGATCCCATCCCA 1503
 DB 1441 AGGCGCTGACCTGCTCCGTTTGAAGTGTGCTGTATCCACAGAGATCCCATCCCA 1500
 QY 1504 TTGACAGACTTGTGTTGAATTCAAAATGAAATTCACCTGCGTCTCAGAGAGCTTCCCTA 1563
 DB 1501 TTGACAGACTTGTGTTGAATTCAAAATGAAATTCACCTGCGTCTCAGAGAGCTTCCCTA 1560

QY 1564 ACCCTGTGAAGACAGACAGCTTTGAGGGGCTCCACCTGCTCTGCTTCCGAC 1623
 DB 1561 ACCCTGTGAAGACAGACAGCTTTGAGGGGCTCCACCTGCTCTGCTTCCGAC 1620
 QY 1624 CCCCCTTCTGTCTCCCTTCTGCTGCTCCATATCTGTTTCTGTCTGCGACCTTCCCT 1683
 DB 1621 CCCCCTTCTGTCTCCCTTCTGCTGCTCCATATCTGTTTCTGTCTGCGACATACACTG 1680
 QY 1684 TCTTC 1688
 DB 1681 TCTCC 1685
 RESULT 7
 ABL68881
 ID ABL68881 standard; DNA; 1763 BP.
 AC ABL68881;
 XX
 DT 15-MAY-2002 (first entry)
 DE
 XX Kidney cancer related gene sequence SEQ ID NO:7218.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 PD 13-DEC-2001.
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 18-SEP-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236039P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 28-SEP-2000; 2000US-236412P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.

PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Edner R, Endreess G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1; SEQ ID 7218; 44dp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX
 SQ Sequence 1763 BP; 381 A; 538 C; 430 G; 414 T; 0 other;

Query Match 63.1%; Score 1625.8; DB 24; Length 1763;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;

QY 10 AGAGATCAGCAGAGTGTGACCAATGATGTCTGTGCTGAGCCCGACAGACTCTGG 69
 DB 1 AGAGATCAGCAGAGTGTGACCAATGATGTCTGTGCTGAGCCCGACAGACTCTGG 60
 QY 70 GTGATGTCTGTGAAATCTTCCAGAGCGGCTCCCTGTCTCAATCTGTCTGTGCTGATCA 129
 DB 61 GTGATGTCTGTGAAATCTTCCAGAGCGGCTCCCTGTCTCAATCTGTCTGTGCTGATCA 120
 QY 130 AGGAGTTTCAAGTCTTCACTGTGACAGGAGTGTGCTGAAAGCCCTCCAGAGTTCCCT 189
 DB 121 AGGAGTTTCAAGTCTTCACTGTGACAGGAGTGTGCTGAAAGCCCTCCAGAGTTCCCT 180
 QY 190 GCCCTCCCTCCCACTGAGCTCTTCCGAGACATCCAGAGTCTCCAAACAGACAGAGACTAC 249
 DB 181 GCCCTCCCTCCCACTGAGCTCTTCCGAGACATCCAGAGTCTCCAAACAGACAGAGACTAC 240
 QY 250 AACGGAATTCAGAAATGGGTGAGAGACATTCCTCAAGTGTCTCTCAATGGGCTATGGGAG 309
 DB 241 AACGGAATTCAGAAATGGGTGAGAGACATTCCTCAAGTGTCTCTCAATGGGCTATGGGAG 300
 QY 310 GCAAGTTCTGTCTCAAGCTCTATGACCTGTGATATATGAAGGTGATTCTGGGGAGATCAG 369
 DB 301 GCAAGTTCTGTCTCAAGCTCTATGACCTGTGATATATGAAGGTGATTCTGGGGAGATCAG 360
 QY 370 ACCGGAATTCAGAAATGGGTGAGAGACATTCCTCAAGTGTCTCTCAATGGGCTATGGGAG 429
 DB 361 ACCGGAATTCAGAAATGGGTGAGAGACATTCCTCAAGTGTCTCTCAATGGGCTATGGGAG 420
 QY 430 TGTGAAATGGGAGACATGATGTTCCAGATCGAAGGATGTGACCCCGAGCTTCAATATG 489

Db 421 TGTGAATGGGAGAGATGGTTCCAGATGACGAGATGCTGACCCGACTTTCACATAG 480
 Qy 490 ACATCTGAAGCCCTATGATGGGGCTATGAGCAGACTCTGATGAGATGCTGAGCAAAAT 549
 Db 481 ACATCTGAAGCCCTATGATGGGGCTATGAGCAGACTCTGATGAGATGCTGAGCAAAAT 540
 Qy 550 GGGAGAGAGCTCTTGGCCAGAGATCCCTCTGAGAGGTCTTTCAGCAGATCTCTTATGA 609
 Db 541 GGGAGAGAGCTCTTGGCCAGAGATCCCTCTGAGAGGTCTTTCAGCAGATCTCTTATGA 600
 Qy 610 CCTGAGACCATCATGAGATGATGCTTACGATCAGAGGAGCAGATCCAGTGAAGAGAG 669
 Db 601 CCTGAGACCATCATGAGATGATGCTTACGATCAGAGGAGCAGATCCAGTGAAGAGAG 660
 Qy 670 ATTCTAGTCTTACATATACAGGACCATTAAGTACCTGAAACAACCTGTTTTTCCGATGA 729
 Db 661 ATTCTAGTCTTACATATACAGGACCATTAAGTACCTGAAACAACCTGTTTTTCCGATGA 720
 Qy 730 GGAATGCTTTTACCGAGATGACACCATCTACAGCTGACCTTGTGCGCCCTGACAC 789
 Db 721 GGAATGCTTTTACCGAGATGACACCATCTACAGCTGACCTTGTGCGCCCTGACAC 780
 Qy 790 ACCGCGCTGCGCAGCTGGCCATCAGACAC---AGACGAGATGATCACTGAGAGAG 846
 Db 781 ACCGCGCTGCGCAGCTGGCCATCAGACAC---AGACGAGATGATCACTGAGAGAG 840
 Qy 847 CTCAACTACAGAGAGAGGGAGCTGAGAGATCAAGAGAGAGAGCATTTGATTTTC 906
 Db 841 CTCAACTACAGAGAGAGGGAGCTGAGAGATCAAGAGAGAGAGCATTTGATTTTC 900
 Qy 907 TGGATATCTCTCTTTGGCCAAAATGAGAGATGAGAGATCTTTGTACAGACAGACCTCC 966
 Db 901 TGGATATCTCTCTTTGGCCAAAATGAGAGATGAGAGATCTTTGTACAGACAGACCTCC 960
 Qy 967 GTCCTGAGGTGACAGTTCATGTTGAGGGCCACGACACACAGCCAGTGGAGATCTCCT 1026
 Db 961 GTCCTGAGGTGACAGTTCATGTTGAGGGCCACGACACACAGCCAGTGGAGATCTCCT 1020
 Qy 1027 GGATCTCTATGCTTGGCCACACACCCAGAGATCAGAGAGGTGCGGAGAGAGATCC 1086
 Db 1021 GGATCTCTATGCTTGGCCACACACCCAGAGATCAGAGAGGTGCGGAGAGAGATCC 1080
 Qy 1087 ACAGCTCTCTGAGTATGAGAGCTTCATCACTTGAACAACCTTGAACAGATGCTTACA 1146
 Db 1081 ACAGCTCTCTGAGTATGAGAGCTTCATCACTTGAACAACCTTGAACAGATGCTTACA 1140
 Qy 1147 CCACCATGTCATTAAGAGAGGACCTGAGGCTTACCAACCGGTGCGCAGGATTTGGCAG 1206
 Db 1141 CCACCATGTCATTAAGAGAGGACCTGAGGCTTACCAACCGGTGCGCAGGATTTGGCAG 1200
 Qy 1207 AGCTCAGACATCCCGTACCTTCCTGATGAGGGGCTCTTGCCCAAGATCATGCTCC 1266
 Db 1201 AGCTCAGACATCCCGTACCTTCCTGATGAGGGGCTCTTGCCCAAGATCATGCTCC 1260
 Qy 1267 TCCTCTCATTTATGAGCTTACCAACAACCAAAAGTGTGGCCCAACCCAGAGGTGTTG 1326
 Db 1261 TCCTCTCATTTATGAGCTTACCAACAACCAAAAGTGTGGCCCAACCCAGAGGTGTTG 1320
 Qy 1327 ACCCTTTCGTTTTCACCGGTTTCTGCTACACAGCCAGCTTTCCTGCTTCTCAG 1386
 Db 1321 ACCCTTTCGTTTTCACCGGTTTCTGCTACACAGCCAGCTTTCCTGCTTCTCAG 1380
 Qy 1387 GAGATCAAGAGACTGATGAGAGAAACAATTTGCAAGAAAGAGCTGAGAGGTGSC---A 1443
 Db 1381 GAGATCAAGAGACTGATGAGAGAAACAATTTGCAAGAAAGAGCTGAGAGGTGSC---A 1440
 Qy 1444 CGGCGCTGAGCCTGCTCGCTTTCAGCTGCTGATCCACAGAGATCCCATCCCA 1503
 Db 1441 AGGCGCTGAGCCTGCTCGCTTTCAGCTGCTGATCCACAGAGATCCCATCCCA 1500
 Qy 1504 TTGACAGACTTGTGTAATTCAAAAATGGAATTCACCTGCTCAGAGAGCTCCCTA 1563
 Db 1501 TTGACAGACTTGTGTAATTCAAAAATGGAATTCACCTGCTCAGAGAGCTCCCTA 1560

Qy 1564 ACCCTTTCAGAGCAAGAGACAGCTTTGAGGGCTCCACTGCGCTCTTCTTCTGAC 1623
 Db 1561 ACCCTTTCAGAGCAAGAGACAGCTTTGAGGGCTCCACTGCGCTCTTCTTCTGAC 1620
 Qy 1624 CCGCGCTGCTGCTCCCTCTGCTGAGCCCAATCTGTTTCTGTCGACCATTCCT 1683
 Db 1621 CCGCGCTGCTGCTCCCTCTGCTGAGCCCAATCTGTTTCTGTCGACCATTCCTG 1680
 Qy 1684 TCTTC 1688
 Db 1681 TCTTC 1685
 RESULT 8
 ABK50070
 ID ABK50070 standard; DNA; 1872 BP.
 XX
 AC ABK50070;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE DNA encoding human Cyp 4A22 protein.
 XX
 KW Human; Cyp 4A14; hypertension; Cyp 4A11; testosterone inhibitor; obesity;
 KW lipid metabolism disease; pancreatic dysfunction; type II diabetes;
 KW cardiovascular disease; Cyp 4A22; gene; ds.
 OS Chimeric - Homo sapiens.
 OS Chimeric - Synthetic.
 XX
 FH Key
 FT CDS
 FT 313..1872
 FT /tag= a
 FT /product= "Human Cyp 4A22 protein"
 XX
 PN W0200217856-A2.
 XX
 PD 07-MAR-2002.
 XX
 PD 29-AUG-2001; 2001MO-US26914.
 XX
 PF 29-AUG-2000; 2000US-228947P.
 XX
 PR 29-AUG-2000; 2000US-228947P.
 XX
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Capdevilla J, Waterman M, Holla V;
 XX
 DR WPI; 2002-382928/41.
 DR P-PSDB; A0079995.
 XX
 PT Treating hypertension in an individual by inhibiting testosterone
 PT activity; enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11
 PT activity in the individual -
 XX
 PS Examples; Page 73-74; 80pp; English.
 XX
 CC The present invention relates to a new method of treating hypertension in
 CC an individual. The method comprises inhibiting testosterone activity,
 CC enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the
 CC individual. Enhancing Cyp 4A14 activity leads to inhibition of
 CC testosterone activity which in turn leads to inhibition of Cyp 4A11
 CC activity. The method of the invention can be used for treating
 CC hypertension in an individual. The molecules of the invention are also
 CC useful for treating disease states associated with lipid metabolism,
 CC pancreatic dysfunction, obesity, type II diabetes, and other
 CC cardiovascular diseases. The present nucleic acid sequence encodes the
 CC human Cyp 4A22 protein of the invention, as described above.
 XX
 SO Sequence 1872 BP; 427 A; 554 C; 481 G; 410 T; 0 other;
 Query Match 59.2%; Score 1525.2; DB 24; Length 1872;
 Best Local Similarity 97.6%; Pred. No. 0;

Matches 1548; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

OY	7	CGCAGAGATCCAGGAGGAGTGTCTGCACCAATGAGTGTCTGTGTGTGAGCCCCAGACAGATCC	66
Db	287	CTCAGAGATCCAGGAGGAGTGTGCACCAATGAGTGTCTGTGTGTGAGCCCCAGACAGATCC	346
OY	67	TGGGATGATGTCTGTGGAATCCCTCAAGGAGGAGCCCTCCGTGTCAATTCGTGCTTGTGTCTG	126
Db	347	TGGGATGATGTCTTCGGGAATCCTTCAGAGTACCTCCCTGTCTCAATTCGTGCTTGTGTCTG	406
OY	127	TCAAGGAGATTCAGGCTCTACCTTGACAAGGAGTGTGTCTCAAGGAGCCCTCCAGAGATTC	186
Db	407	TCAAGGAGAGCTAGGCTCTACCTTGACATAGGAGTGTGTCTCAAGGAGCCCTCCAGAGATTC	466
OY	187	CGTGCCTTCCTCCCATCTGGCTCTTCGGGACATTCAGAGAGCTCCAAAGAGCCAGAGAGC	246
Db	467	CGTGCCTTCCTCCCATCTGGCTCTTCGGGACATTCAGAGAGTTCAAACAGAGCAGAGAGC	526
OY	247	TACAACGGAATTCAGAAATGTGGGTGGAGACATTCCCAAGTGTGTCTCAATTTGGCTAATGG	306
Db	527	TACAACGGAATTCAGGAACGGGTGGAGAGACATTCCAAAGTGTGTCTCAATTTGGGAATGG	586
OY	307	GAGGCAAGATTCATGTCCAGCTCTAATGACCCCTGACCTAATATGAGAGTGAATTCAGGAGAT	366
Db	587	GAGGCAAGATTCATGTCCAGCTCTAATGACCCCTGACCTAATATGAGAGTGAATTCAGGAGAT	646
OY	367	CAGAGCCGAAATCCATGATGTTCTCAACAGATTCCTGTGCTCAATGGAATGGGATACGGCTTC	426
Db	647	CAGAGCCGAAATCCCATATGTTCTCAACAGATTCCTGTGCTCAACAGATTTGGGATACGGCTTC	706
OY	427	TCTGTGTAATGGGAGACATGTGTTCCAGCATCGAGGATGTGACCCCAAGCTTCCACT	486
Db	707	TCTGTGTGAATGGGAGACATGTGTTCCAGCATCGAGGATGTGACCCCAAGCTTCCACA	766
OY	487	ATGACATCTGGAAGCCATATGTGGGGCTCAATGGCAGACCTGTATGAGATGATGTGAGCA	546
Db	767	ATGACATCTGGAAGCCATATATGTGGGGCTCAATGGCAGACCTGTATGAGATGATGTGAGCA	826
OY	547	AATGGGAAGACTCTTGTGGCCAGGAATTCCTCTGAGAGTCTTTCAAGACGCTCTCTTGA	606
Db	827	AATGGGAAGAGTCTTGTGGCCAGGAATTCCTCTGTGAGAGTCTTTCAAGACGCTCTCTTGA	886
OY	607	TGACCTTGGACACATCATGAATGTGCTTCAAGCATCAGGGCAGCATCCAGGTGAGCA	666
Db	887	TGACCTTGGACACCATCATGAAGTGTGCTTCAAGCATCAGGGCAGCATCCAGGTGAGCA	946
OY	667	GGAATTCCTCAATCCCTACATACAGGCAATTAAGAGCTGACAAACCTGTGTTTTTCCCGTG	726
Db	947	GGAATTCCTCAATCCCTACATACAGGCAATTAAGAGCTGACAAACCTGTGTTTTTGTCTGA	1006
OY	727	TGAGGAATGCTTTTACACAGAAATGACACCATCTACAGCTGACCTCTGTGGCCGTGGA	786
Db	1007	TGAGGAATGCTTTTATATAGAAATGACACCATCTACAGCTGACCTCTGTGTGGCCGTGGA	1066
OY	787	CACACCGGCTGTGCAGCTGTGCCCATCAGACACAGACCAAGTATCCAACTTGAGAAAGG	846
Db	1067	CACACCGGCTGTGCAGCTGTGCCCATCAGACACAGACCAAGTATCCAACTTGAGAAAGG	1126
OY	847	CTCAACTATCAGAAAGAGGGGAGGCTGTGAGAAAGTCAAGAGAAAGGCAATTTGATTTTC	906
Db	1127	CTCAACTATCAGAAAGAGGGGAGGCTGTGAGAAAGTCAAGAGAAAGGCACTTTGATTTTC	1186
OY	907	TGATATATCTCTCTTGTGCAAAATATGAGAGATTTGGAGGATCTTGTACAGACAGACCTCC	966
Db	1187	TGATATATCTCTCTTGTGCAAAATATGAGAGATTTGTGTACAGACAGACCTCC	1246
OY	967	GTCGTAGAGTGGACAGTTCAATGTTTGAAGGCAAGACCAACAGCCAGTGGGAATCTTCT	1026
Db	1247	GTCGTAGAGTGGACAGTTCAATGTTTGAAGGCAAGACCAACAGCCAGTGGGAATCTTCT	1306
OY	1027	GGATTCCTCTAATGTCTTGGCCACACACCCCAAGCATCAGAGAGAGTGTCCGGAGAGAGATCC	1086
Db	1307	GGATTCCTCTAATGTCTTGGCCACACACCCCAAGCATCAGAGAGAGTGTCCGGAGAGAGATCC	1366

Qy	1087	ACAGCCCTCCGGGGAATGAGAGCCCTCAATCACTGGAAACCACTGGACCAAGATGCCCTACA	1146
Ds	1367	ATGGCCTCTCTGGGATGATGAGAGCTTCATCACTGGAAACCACTGGACCAAGATGCCCTACA	1426
Qy	1147	CCACCAATGTCATTAAAGAGGCACTGAGGCTCTAACCAACGGGTGCCAGGCAATTGGCCAGAG	1206
Ds	1427	CCACCAATGTCATTAAAGAGGCACTGAGGCTCTAACCAACGGGTGCCAGGCAATTGGAAAGAG	1486
Qy	1207	AGCTCAACACTCCCGTCACCTTCCCTGATGAGGCGCTCTTTGCCCAAGGTAATCAATGCTCC	1266
Ds	1487	AGCTCAACACTCCCGTCACCTTCCCTGATGAGGCGCTCTTTGCCCAAGGTAATCAATGCTCC	1546
Qy	1267	TCCCTCTCATTATATAGCGCTTCAACCAACCCCAAAAGTGTGGCCCAACCCAGAGGTGTTTG	1326
Ds	1547	TCCCTCTCATTATATAGCGCTTCAACCAACCCCAAAAGTGTGGCCCAACCTAGAGGTGTTTG	1606
Qy	1327	ACCCCTTCGGTTTTGCAACGGGGTTCTGCTCAACAAGCCACGGCTTTCCTAGCCCTTCACAG	1386
Ds	1607	ACCCCTTCGGTTTTGCAACGGGGTTCTGCTCAACAAGCCACGGCTTTCCTAGCCCTTCACAG	1666
Qy	1387	GAGGATCAAGGAATGCTGATTTGGGAAAACAATTTGCGATGAACAGAGCTGAAGGTGGCCACGG	1446
Ds	1667	GAGGATCAAGGAATGCTGATTTGGGAAAACAATTTGCGATGAACAGAGCTGAAGGTGGCCACGG	1726
Qy	1447	CCCTGACCCCTGCTCCGCTTTGAGCTGTGCTGTATCCACACAGGATCCCATCCCATTTG	1506
Ds	1727	CCCTGACCCCTGCTCCGCTTTGAGCTGTGCTGTATCCACACAGGATCCCATCCCATTTG	1786
Qy	1507	CACGACTTGTGTGAATCCAAATAAGAAATCAACCTGCTTCAGAGAGCTCCCTTAAC	1566
Ds	1787	CACGACTTGTGTGAATCCAAATAAGAAATCAACCTGCTTCAGAGAGCTCCCTTAAC	1846
Qy	1567	CTTGTGAAGCAAGGACCAAGCTTTGA	1592
Ds	1847	CTTGTGAAGCAAGGACCAAGCTTTGA	1872

ID	ABKS0068	standard; DNA; 2116 BP.
XX	AC	ABKS0068;
XX	XX	
XX	DE	DNA encoding mouse Cyp 4A12 protein.
XX	KM	Mouse; Cyp 4A14; hypertension; Cyp 4A11; testosterone inhibitor; obesity
XX	KW	lipid metabolism disease; pancreatic dysfunction; type II diabetes;
XX	KW	cardiovascular disease; Cyp 4A12; gene; ds.
OS	Chimeric - Mus sp.	
XX	Chimeric - Synthetic.	
FH	Key	Location/Qualifiers
FT	CDS	282..1808
FT	FT	/tag= a
FT	FT	/product= "Mouse Cyp 4A12 protein"
FT	FT	/trans_except= (pos:1044..1046, aa:Gln)
XX	PN	MO200217856-A2.
XX	PD	07-MAR-2002.
PF	PF	29-AUG-2001; 2001WO-US26914.
XX	PR	29-AUG-2000; 2000US-228947P.
XX	PA	(UYVA-) UNIV VANDERBILT.
XX	PI	Capdevila J, Waterman M, Holla V,

DR MPI, 2002-382929/41.
DR P-PSDB; AAU79993.

PT Treating hypertension in an individual by inhibiting testosterone
PT activity, enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11
PT activity in the individual -

XX Claim 5; Page 69-70; 80pp; English.

XX The present invention relates to a new method of treating hypertension in
XX an individual. The method comprises inhibiting testosterone activity,
XX enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the
XX individual. Enhancing Cyp 4A14 activity leads to inhibition of
XX testosterone activity which in turn leads to inhibition of Cyp 4A11
XX activity. The method of the invention can be used for treating
XX hypertension in an individual. The molecules of the invention are also
XX useful for treating disease states associated with lipid metabolism,
XX pancreatic dysfunction, obesity, type II diabetes, and other
XX cardiovascular diseases. The present nucleic acid sequence encodes the
XX mouse Cyp 4A12 protein of the invention, as described above.

Sequence 2116 BP; 517 A; 568 C; 441 G; 590 T; 0 other;

Query Match 42.0%; Score 1082.8; DB 24; Length 2116;
Best Local Similarity 76.9%; Pred. No. 5.4e-239;
Matches 1375; Conservative 0; Mismatches 402; Indels 11; Gaps 4;

QY 11 GAGATCAGACAGGCTGACACATGAGTCTCTGCTGAGCCGACAGACTCTGGG 70
DB 260 GTATCTCAGAGCTGTTGATCATGATGCTCTGCTCTCATGCTCTCATGATTTCCAGG 319
QY 71 TGATGCTCTGGAATCTTCCAGCGGCTCCCTGCTCATTTCTGCTGCTGATCAA 130
DB 320 AAGCATCTGAGTACCTTCAAGTAGCTCTGCTGCTGCTGCTGCTGCTGCTCAA 379
QY 131 GGGATTCAGCTCTACCTGACAGAGGAGTGGCTCTCAAGACCTCCAGAGTTCCCGTG 190
DB 380 GAGAGCCAGCTCTACCTGACAGAGGAGTGGCTCTCAAGAGTCTGAGAGTTCCATC 439
QY 191 CCCCTCCCTCCAGCTCTTCCGGGCAATCCAGAGGCTCCAGAGGAGGAGTCA 250
DB 440 CCCACCTTCTCAGCTGCTTTTGGACACA-----AGATCTTAAGAGACAGAGCTTCA 493
QY 251 ACCGATTCAGAAATGGGTGAGACATTCCTCAAGTCCCTGCTCATTTGGCTATGGGAGG 310
DB 494 AGATATCTTAAGTATGAGTAAAGATTTCCAAAGTCCCTGCTCAAGTGGCTCTGGGAGG 553
QY 311 CAAGATTCGTGTCAGCTCTATGACCTGACTATATGAGTATTTGGGGAGATCAGA 370
DB 554 CAAGATTCGTGTCAGCTCTATGACCTGACTATATGAGTATTTGGGGAGATCAGA 613
QY 371 CCCGAATCCAGTGTCTTCAAGATTCCTGGCTCCAGATTTGGGATAGGCTGGCTCT 430
DB 614 CCCAAAGCTTAATGTTCTCAAGATTTCTAGCTCCCTGAGTGGGCTGGTGTCTTAT 673
QY 431 GTTGAATGGGAGACATGTTTCCAGATTCAGAGATGCTGACCCAGCTTCCATATGA 490
DB 674 GCTGATGAGACAGATGTTTCAAGACCGAGAGATGATGACCCAGCTTCCATATGA 733
QY 491 CATCTGAAGCCCTATGTTGGGCTCATAGGACACTCTGATGAGTATGAGCAAAATG 550
DB 734 CATCTGAAGCCCTATGTTGGGCTCATAGGACACTCTGATGAGTATGAGCAAAATG 793
QY 551 GGAAGACTCCTTGGGAGAGATTCCTCTGAGAGTCTTTGAGAGAGCTCTCTGATAG 610
DB 794 GGAAGACTTGTGGGAGAGATTCCTCTGAGAGTCTTTGAGAGAGCTCTCTGATAG 853
QY 611 CTTGAGACCATCATGAGTGTGCTTCAAGCATTCAGAGGAGAGATTCAGTGTGACAGAA 670
DB 854 CTTGAGACCATCATGAGTGTGCTTCAAGCATTCAGAGGAGAGTGTGAGTGTGACAGAA 913
QY 671 TTCTCAGTCTCATATACAGGCAATTAAGTACCTGAACCACTGTTTTCCTGTTAG 730

DB 914 ATACAGTCTATATTCAGGACGTGAGACCTGAACGATCTGTTTTTCCCTGTGCG 973
QY 731 GAATGCTTTTACCAAGATGACACCATCTACAGCTTGAACCTTGTGCGCGCTGAGACA 790
DB 974 GAACATCTTTCACCTGATGATCATCTACAGAGTCTCTATGAGCTGCAAGGCTTAA 1033
QY 791 CCGGCTCTGACACTGGCCCATGACAGACAGACCAAGTATCCACTGAGGAGGCTCA 850
DB 1034 CAGTGCCTGAACTTGGCCCATATATCAACAGACCAAGTATCAATCAAGAGAGATCA 1093
QY 851 ACTACAGAGAGAGGAGCTGAGAGATCAAGAGAGAGGAGGATTTGATTTCTGGA 910
DB 1094 ACTTCAAGATGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1153
QY 911 TATCTCTCTTGGCCAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 970
DB 1154 CATCTCTCTTGGCCAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1213
QY 971 TGAAGTGAACAGTTGATGTTTGAAGGCGACAGACAGAGCGGAGGAGATCTCTGAT 1030
DB 1214 TGAAGTGAATCTTCAATGTTTGAAGGCGACAGACAGAGCTATGATCTCTGAT 1273
QY 1031 CCTTATGCTCTGAGCCACACACCCAAAGATCAAGAGAGTGGCCGAGAGAGATCCAG 1090
DB 1274 CTTCTATGCTTTTGGCCCAATCTTGAACATCAACAGAGATGAGAGAGAGATCCAAAG 1333
QY 1091 CTTCTGAGTATGAGAGCTTCAATCACTGGAACCACTGGAACCAATGAGAGAGAGAG 1150
DB 1334 TCTCTGAGAGATGAGAGCTTCAATCACTGGAACCACTGGAACCAATGAGAGAGAGAG 1393
QY 1151 CATGTCATTAAG 1210
DB 1394 CATGTCATTAAG 1453
QY 1211 CAGCATCCCGTCACTTCCCTGATGAGGAGCTCTTCCCAAGAGATCATGATCTCTCT 1270
DB 1454 CAGTCATCTGTCACCTTTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
QY 1271 CTCATTTATGAGCTTTCACCAACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1330
DB 1514 CTCCTTTATGAGCTTTCATCAACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1573
QY 1331 TTTCCGTTTTCAGAGGAGTCTGTCACACAGGAGCTTCTGCGCTTCTCAGAGAG 1390
DB 1574 TTTCCGTTTTCAGAGGAGTCTTCCCGGACAGGAGCTCATCTCTGCTTCTCAGAGAG 1633
QY 1391 ATCAAGAGATGCAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1450
DB 1634 AGCAAGAGATGCAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1693
QY 1451 GACCTGCTCGCTTTCAGAGTGTGCTGATTCACAGAGATCCCATTCCTCATTTGACG 1510
DB 1694 GACCTGCTCGCTTTCAGAGTGTGCTGATTCACAGAGATCCCATTCCTCATTTGACG 1753
QY 1511 ACTTGTGTTGAATTCAG 1570
DB 1754 AATGTGTGTAAGTTCAG 1813
QY 1571 TGAAG 1630
DB 1814 AGAGAGAGAG--AGAGCTCAAGATGAGTGTGCTGATTCATGCTTCTCTGACTTACT 1871
QY 1631 TCTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1689
DB 1872 CTTTTCCTCAATCTTTCATGCTGATCATGCTCATCTTCTTCTTCTTCTTCTTCTTCT 1931
QY 1690 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1747
DB 1932 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1991
QY 1748 CCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
DB 1992 CCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2039

RESULT 10
ABK63715

ID ABK63715 standard; cDNA; 2462 BP.

AC ABK63715;

DT 18-JUN-2002 (first entry)

DB Rat sequence differentially expressed in response to a hepatotoxin #1622.

Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening; differential expression; centrilobular necrosis; steatosis.

Rattus norvegicus.

PN WO200210453-A2

PD 07-FEB-2002.

PF 30-JUL-2001; 2001WO-US23872

PR 31-JUL-2000; 2000US-222040P

PR 11-MAY-2001; 2001US-290029P

PR 22-MAY-2001; 2001US-292336P

PR 13-JUN-2001; 2001US-297457P

PR 09-JUL-2001; 2001US-303459P

PA (GENE-) GENE LOGIC INC.

PI Mendrick D, Porter MW, Johnson

WPI; 2002-241625/29.

PT Predicting toxic effects of

cells exposed to the to

XX

XX

CC compounds or the progression of these toxic effects by determining the

toxin and comparing these to gene expression

effect of a compound or progression of a toxic effect, preferably the hepatotoxicity of a compound, comprising detecting the level of expression in a tissue or cell sample exposed to the compound of two or more genes listed in the specification, where differential expression of the genes is indicative of at least one toxic effect or progression. The method can also be used to identify an agent which modulates the toxic response and predict cellular pathways that a compound modulates in a cell. The methods utilise a set of at least two probes (on a solid support in kit form), where each of the probes comprises a sequence that specifically hybridises to a gene listed in the specification, a computer system comprising a database containing information identifying the expression level in a tissue or cell sample exposed to a hepatotoxin of a set of genes comprising at least two genes listed in the specification, and a user interface to view the information used to present information identifying the expression level in a tissue or cell of at least one gene listed in the specification. The method is useful for elucidating global changes in gene expression and for identifying toxicity markers in tissues or cell exposed to a known toxin. The genes may be used as toxicity markers in drug screening and toxicity assays. The genes and gene expression information may be used as diagnostic markers for the prediction or identification of the physiological state of tissue or cell sample that has been exposed to a compound or agent. Hepatotoxicity

CC Sequence 2462 BP; 676 A; 643 C; 507 G; 636 T; 0 other;
XQ
CC is characterised by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent.
XQ

Sequence 2462 BP; 676 A; 643 C; 507 G; 636 T; 0 other

Query Match	39.7%	Score 1022.4	DB 24	Length 2462
Best Local Similarity	79.6%	Pred. No. 4.5e-225		
Matches 1222, Conservative	0	Mismatches 311	Indels 3	Gaps 1

QY	25	GCTGCACATGAAAGTCTCTGTGCTGACGCCCAAGACATCTCGGGTGAATGTCTGGAA	84
Db	5	GCTGCACATGAAAGTCTCTGTGACCTGAGCTCAACCCGTTCAAGGGACAGATCTTGGCT	64
QY	85	TCTCTCAAGCGGCGCTCCCTGCTCATTTCTGTCTGTCTGATCAAGGAGTTCAAGCTCT	144
Db	65	TCTCTCAAGTGGCTCCGCTGCTGGTGTGATCTGTCTGTCTGATCAAGGAGTTCAAGCTCT	124
QY	145	ACCTGCACAGGCAAGTGGTGTCTCAAAAGCCCTTCAGACAGTTCCCGTGCCTCTCCCACT	204
Db	125	ACCTGCACAAAGGCATATGGCTACTCAAGGCTTTTCAGACAGTTCCCACTCTCCCTTCCACT	184
QY	205	GGGCTCTGGGGCAATCCAGGAGGCTCCAAAGGACAGGAGCTACAAGGATTCAGAAAT	264
Db	185	GGTTCTTTGGGCAACA--AGCAATTTCAAGGTGACAAAGAACTACACGAAATTAATACAT	244
QY	265	GGGTGGAGACATTTCCCAAGTGCCTGTCTCTCATTTGGCTATGGGAGGCAAAAGTTCTGTCC	324
Db	242	GTGTGGAGAAATTTCCCAAGTGCCTTTCTCGATGGTTCTGGGGAAGCAAAAGCTACTTAA	304
QY	325	AGCTCTATGACCTGACTATATGAAGTGAATTTGGGAGATCAAGCCGAAATCCCATG	384
Db	302	TTGTCTATGACCCCTGACTACATGAAGGGAATTTCTGGGCGAATCAAGTCCAAAGGCCAATG	364
QY	385	GTTCCTACAGATTCCTGGCTCCATGGAATGGGTGAGGGCTGTCCGTGGAATGGGCAAG	444
Db	362	GCGCTACAGATTTCTAGCTCTTGGAATGGATATGATTTGCTCTTGTCTGAATGACACAC	424
QY	445	CATGTTCCAGACATCGACGAGATGCTGACCCGAGCCTTCCACTATGACATCTGAGACCTT	504
Db	422	CGTGTCTTCAGACACGGGGAAATGCTAACCCAGCCTTCCACTATGACATTTGGAACCTT	484
QY	505	ATGTGGGGCTCATGGCAAGCTCTGTAGAGTATGCTGGACAAATGGGAAGAGCTCTTG	564
Db	482	ATGTAAATAAACATGGCTGACTCCATTCGACTGATCTGACAAATGGGAACAGCTGGCAG	544
QY	565	GCCAGAGTTCCCTCTGGAAGTCTTTTACAGACGTCTCTGATGAGCCCTGGACACACATCA	624
Db	542	GTCAAGACTCCTCATATGAAATCTTTCAACATATCTCTTAATGAGCCCTAGACACGTGTCA	604
QY	625	TGAAGTGTGCTTCAGCCATCGAGGGACAGATTCAGGTGGACAGAAATTCAGTCCCTTACA	684
Db	602	TGAAGTGTGCTTCAGCCACAAATGGCAGTGTTCAGGTGGATGGAAATTAAGAGAGTATA	664
QY	685	TACAGGCTAATTAGTACCTGAAACAACCTGTTTTTCCCGTGTAGGAATGCTTTCAACC	744
Db	662	TTCAGGCTCATTTGGGAACCTTGAATAGACCTTTCACTCCCGTGTAGGAACATCTTTCATC	724
QY	745	AGAAATGACACATTAAGCCTGACCTGTGCTGGCGCGCTGACACACCGCGCTGCGACG	804
Db	722	AGAAATGATACCATTAATATATTTTCTTCOAATGGCCACTGTTCACACGCTGTGTACAC	784
QY	805	TGGGCCATTCAGACACACAGCCAAAGTGAATCCAATGAGGAAGGCTCAATTCAGAAAGAG	864
Db	782	TTTGCCCATATGACACACAGATGTGTGTGATCAAGCTAAGGAAGGATCAGCTGACGAATGCG	844
QY	865	GGGAGCTTGAAGAGATCAAGAGAAAGGCAATTTGGATTTTCTGGAATATCTTCTCTTG	924
Db	842	GAGAGCTGGAAGAGGTCAAGAAAGAAAGACGTTTGGATTTTCTGGAACATCTCTTACTTG	904
QY	925	CCAAATATGGAATATGGAGACATTTGTTCAGACAAAGGACCTCGTGTGAGGTGGACACAT	984
Db	902	CCGAATATGGAATATGGGACACGCTTGTGTGACAAAGACCTTACGTGTGAGGTGGACACAT	964

QY 985 TCATGTTGAGGGCCACGACCAACGACGAGGATCTCTGATCTCTGATCTCTG 1044
 DB 962 TTAATGTTCCAGGGGCTCATGACACCAAGCCAGGAGTCTCTGATCTCTGATCTCTG 1021
 QY 1045 CCACACACCCCAAGCATCTGAGAGAGGTCGGGAGAGATCCACAGCTCTCTGAGGATG 1104
 DB 1022 CCACACACCCCAAGCATCTGAGAGAGGTCGGGAGAGATCCACAGCTCTCTGAGGATG 1081
 QY 1105 GAGGCTTCATCACTCTGAAACCACTCTGAGACCACTCTGACACCATCTGATTAAG 1164
 DB 1082 GGTCTCTCATTAAGTCTGAGATCACTCTGAGACCACTCTGACACCATCTGATTAAG 1141
 QY 1165 AGGCACTGAGGCTCTGACCAAGGTCGAGGATCTGAGAGAGCTGACATCTCTGCA 1224
 DB 1142 AGGCTCTGAGGCTCTGACCAAGGTCGAGGATCTGAGAGAGCTGACATCTCTGCA 1201
 QY 1225 CTTCTCTGATGAGGCTCTCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGG 1284
 DB 1202 CTTCTCTGATGAGGCTCTCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGG 1261
 QY 1285 TTCAACCAACACCAAGGTCGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGG 1344
 DB 1262 TCACACCAACACCAAGGTCGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGG 1321
 QY 1345 CGGCTCTGCTCAACACAGGCAAGCTCTCTGAGGATCTGAGGATCTGAGGATCTGAGG 1404
 DB 1322 CAGACTCTCTGCTCAACACAGGCAAGCTCTCTGAGGATCTGAGGATCTGAGGATCTGAGG 1381
 QY 1405 TTGAGGAAACATTTGTCATGAGAGGTCGAGGATCTGAGGATCTGAGGATCTGAGG 1464
 DB 1382 TTGAGGAAACATTTGTCATGAGAGGTCGAGGATCTGAGGATCTGAGGATCTGAGG 1441
 QY 1465 TTGAGGTCGCTGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGG 1524
 DB 1442 TTGAGGTCGCTGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGG 1501
 QY 1525 CCACCAATGAGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGG 1560
 DB 1502 CCACCAATGAGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGG 1537

RESULT 11

ABK50071

ABK50071 standard; DNA; 21990 BP.

XX

15-JUL-2002 (first entry)

Human CYP 4A11 genomic DNA sequence.

Human; CYP 4A11; hypertension; CYP 4A11; testosterone inhibitor; obesity; lipid metabolism disease; pancreatic dysfunction; type II diabetes; cardiovascular disease; ds.

Chimeric - Homo sapiens.

Chimeric - Synthetic.

W0200217856-A2.

07-MAR-2002.

29-AUG-2001; 2001MO-US26914.

29-AUG-2000; 2000US-228947P.

(UYVA-) UNIV VANDERBILT.

Capdevilla J, Waterman M, Holla V;

WPI; 2002-382929/41.

XX

PT Treating hypertension in an individual by inhibiting testosterone
 PT activity, enhancing CYP 4A14 activity, or by inhibiting CYP 4A11
 PT activity in the individual -
 XX
 PS Claim 11; Page 74-80; 80pp; English.
 CC The present invention relates to a new method of treating hypertension in
 CC an individual. The method comprises inhibiting testosterone activity in
 CC enhancing CYP 4A14 activity, or inhibiting CYP 4A11 activity in the
 CC individual. Enhancing CYP 4A14 activity leads to inhibition of
 CC testosterone activity which in turn leads to inhibition of CYP 4A11
 CC activity. The method of the invention can be used for treating
 CC hypertension in an individual. The molecules of the invention are also
 CC useful for treating disease states associated with lipid metabolism,
 CC pancreatic dysfunction, obesity, type II diabetes, and other
 CC cardiovascular diseases. The present nucleic acid sequence represents
 CC the human CYP 4A11 genomic DNA sequence of the invention, as described
 CC above.
 CC
 SQ Sequence 21990 BP; 1784 A; 2031 C; 1682 G; 1699 T; 14594 other;

Query Match 37.6%; Score 968.4; DB 24; Length 21990;
 Best Local Similarity 98.0%; Pred. No. 2.6e-212;
 Matches 1002; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

QY 1381 TCTCAGAGATCAAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1440
 DB 13759 TCTCAGAGATCAAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 13818
 QY 1441 CCAAGGCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1500
 DB 13819 CCAAGGCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 13878
 QY 1501 CCAATGAGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1560
 DB 13879 CCAATGAGATCTGATGAGGATCAATTTGCAAGAGAGGATG 13938
 QY 1561 CTAACCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1620
 DB 13939 CTAACCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 13998
 QY 1621 GACCCGCTCTGATGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1680
 DB 13999 GACCCGCTCTGATGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 14058
 QY 1681 CTTCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1740
 DB 14059 CTTCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 14118
 QY 1741 CCAAGGCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1800
 DB 14119 CCAAGGCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 14178
 QY 1801 AAGGCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1860
 DB 14179 AAGGCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 14228
 QY 1861 TCTGATCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1920
 DB 14239 TCTGATCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 14298
 QY 1921 TCTGATCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 1980
 DB 14299 TCTGATCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 14358
 QY 1981 CTTCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 2040
 DB 14359 CTTCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 14417
 QY 2041 GCTCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 2100
 DB 14418 GCTCTCTGAGGATCTGATGAGGATCAATTTGCAAGAGAGGATG 14477

QY 2101 CATTATCTCAAGTACAGAGCAGATGCTAAACCTGCACAGCTGTAATTTGGCTTAT 2160
 DB 14478 CATTATCTCAAGTACAGAGCAGATGCTAAACCTGCACAGCTGTAATTTGGCTTAT 14537
 QY 2161 GCAACATGCTCTTTCAGACCCCACTTCACACCTGTTTCTTTGGTATCACC 2220
 DB 14538 GCAACATGCTCTTTCAGACCCCACTTCACACCTGTTTCTTTGGTATCACC 14594
 QY 2221 ATAAATATCTGCACTTTCAGAGCTGCGGGCTTCACAGCTTCATCTTACCTTTGGCG 2280
 DB 14595 ATAAATATCTGCACTTTCAGAGCTGCGGGCTTCACAGCTTCATCTTACCTTTGGCG 14654
 QY 2281 CCTGAGACCACTTCTCTCTCAACCTGTTTCTCACTGCTTGACTGCGCGACTT 2340
 DB 14655 CCTGAGACCACTTCTCTCTCAACCTGTTTCTCACTGCTTGACTGCGCGACTT 14714
 QY 2341 TGTCAACCCCAAGACCTGCTGTTGGGTGGAACACCCCAATCCCTGATCTCCACCA 2400
 DB 14715 TGTCAACCCCAAGACCTGCTGTTGGGTGGAACACCCCAATCCCTGATCTCCACCA 14774
 QY 2401 CC 2402
 DB 14775 CC 14776

RESULT 12

ABK50067 standard; DNA; 4123 BP.

ABK50067;

15-JUL-2002 (first entry)

DNA encoding mouse Cyp 4A14 protein.

Mouse; Cyp 4A14; hypertension; Cyp 4A11; testosterone inhibitor; obesity;

lipid metabolism disease; pancreatic dysfunction; type II diabetes;

cardiovascular disease; gene; ds.

Chimeric - Mus sp.

Chimeric - Synthetic.

Location/Qualifiers

/*tag= a

/product= "Mouse Cyp 4A14 protein"

/transl_except= (pos:1649..1651, aa:Val)

/transl_except= (pos:1832..1835, aa:Asp)

/note= "This codon has an apparent 1 nucleotide

insertion which alters the reading frame"

MO200217856-A2.

07-MAR-2002.

29-AUG-2001; 2001WO-US26914.

29-AUG-2000; 2000US-228947P.

(UTVA-) UNIV VANDERBILT.

Capdevila J, Waterman M, Holla V;

WPI; 2002-38929/41.

P-PSDB; AAU79992.

Treating hypertension in an individual by inhibiting testosterone

activity, enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11

activity in the individual -

Claim 1; Page 65-66; 80pp; English.

The present invention relates to a new method of treating hypertension in

CC an individual. The method comprises inhibiting testosterone activity,
 CC enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the
 CC individual. Enhancing Cyp 4A14 activity leads to inhibition of Cyp 4A11
 CC testosterone activity which in turn leads to inhibition of Cyp 4A11
 CC activity. The method of the invention can be used for treating
 CC hypertension in an individual. The molecules of the invention are also
 CC useful for treating disease states associated with lipid metabolism,
 CC pancreatic dysfunction, obesity, type II diabetes, and other
 CC cardiovascular diseases. The present nucleic acid sequence encodes the
 CC mouse Cyp 4A14 protein of the invention, as described above.
 CC XX

Sequence 4123 BP; 1130 A; 860 C; 867 G; 1265 T; 1 other;

Query Match 37.5%; Score 966; DB 24; Length 4123;

Best Local Similarity 73.5%; Pred. No. 5.1e-212;

Matches 1298; Conservative 0; Mismatches 440; Indels 28; Gaps 4;

30 ACCATGATGCTCTGCTGCTGAGCCCAAGACCTCTGCTGATCTCTGATCTCTG 89

DB 1634 ACCATGATGCTCTGCTGCTGAGCCCAAGACCTCTGCTGATCTCTGATCTCTG 1693

90 CAAGCGCT 149

DB 1694 CAATGGCT 1753

150 CACAGGAGTGGCT 209

DB 1754 CCAAGGAGTGGCT 1813

210 TTGGGGGACATCCAGAGCTCTCAAGACCAAGAGCTTCAACGGATTGAAATGGGTG 269

DB 1814 TTGGGGGACATCCAGAGCTCTCAAGACCAAGAGCTTCAACGGATTGAAATGGGTG 1865

270 GAGACATTCCTCAAGGCT 329

DB 1866 GAGACATTCCTCAAGGCT 1925

330 TATGACCT 389

DB 1926 TATGACCT 1985

390 TACAGATTCCT 449

DB 1986 TATCAATTCCT 2045

450 TTCCAGATTCAGAGGCT 509

DB 2046 TTCCAGATTCAGAGGCT 2105

510 GGGCTCATGAGGCT 569

DB 2106 GGGCTCATGAGGCT 2165

570 GATTCCT 629

DB 2166 GATTCCT 2225

630 TGTGCTTCAAGGCT 689

DB 2226 TGTGCTTCAAGGCT 2285

690 GCAATTAAGTCT 749

DB 2286 GCAATTAAGTCT 2345

750 GACACCATCTCAAGCT 809

DB 2346 GACACCATCTCAAGCT 2405

810 CATCAGACCAAGAGTCT 869

DB 2406 CATCAGACCAAGAGTCT 2465

```

OY 870 CTGGAAGATCAAGAGAGAGGCAATTGATTTCTGATATCTCTCTTGGCCAA 929
DB 2466 CTGCAAGAGGCCAGAGAGAGAGACACTTGGATTCTTGGACATCTCTTGTGGCACA 2525
OY 990 ATGAGAGATGGAGAGATCTTTCAGACAGAGACCTCCGATGAGAGTGAACAGTTGATG 989
DB 2526 ATGAGAGATGGAGAGACCTTGTCTGTATGAGACCTGGCTGACAGAGGTGAACATTCATG 2585
OY 990 TTGAGGAGCCAGACAGACCAAGCCAGTGGATCTCTGATCTCTATGCTCTGGCCACA 1049
DB 2586 TTGAGGAGCCAGACAGACCAAGCCAGTGGATCTCTGATCTCTATGCTCTGGCCACA 2645
OY 1050 CACCCCAAGCATCAGAGAGAGTGGCCGAGAGAGATCCAGCCTCTGGGTATGAGGC 1109
DB 2646 CACCCCAAGCATCAGAGAGAGTGGCCGAGAGAGATCCAGCCTCTGGGTATGAGGC 2705
OY 1110 TCCATCAGCTGAGACCACTGAGACCAAGCCCTACACACATGTCATTAAGAGGGA 1169
DB 2706 TCTGTACATGGAGACCATCTGAGCCAGATGCCCTACACACATGTCATTAAGAGGGA 2765
OY 1170 CTGAGGCTCTACCCACCGGTGACAGGCAATTGGCAGAGAGCTCAGCCTCCGTCACCTTC 1229
DB 2766 CTGAGGCTCTACCCACCGGTGACAGGCAATTGGCAGAGAGCTCAGCCTCCGTCACCTTC 2825
OY 1230 CCTGATGGGCGCTCTCTGGCCAAAGATATGATGCTCTCTCATTTATGGCCTTCA 1289
DB 2826 CCAATGAGAGCGCTCCATCCAAAGGATATCAGCAGCAATTTATGAGCCTTACAT 2885
OY 1290 CACACCCCAAGAGTGGGCGCAACCGAGAGTGTGACCTTCCGTTTGGACCGGCT 1349
DB 2886 CATAACCCAGTGTCTGGGCGCAACCGAGAGTGTGACCTTCCGTTTGGACCGGCT 2945
OY 1350 TCTGCTCAACAGCAGCAGCTTCTGCTCCCTTCTCAGAGAGATCAAGAGATGATGGG 1409
DB 2946 TCTTCTCAGCAGAGCAGCTTATCTGCAATCTCAGAGAGATCAAGAGATGATGGG 3005
OY 1410 AAACAAATTTGCATGAAAGAGCTGAAAGTGGCAGCGCCCTGACCTGCTCGCTTTGAG 1469
DB 3006 AAACAAATTTGCATGAAAGAGCTGAAAGTGGCAGCGCCCTGACCTGCTCGCTTTGAG 3065
OY 1470 CTGCTGCTGATCCCAAGCAGATCCCAATCCCAATGACAGATTTGTTGAATCCAAA 1529
DB 3066 TTGCTGCGAGATCCCAAGAGATCCCAATCCCAATGACAGATTTGTTGAATCCAAA 3125
OY 1530 AATGAAATCCATGCTGCTCAGAGAGAGCTCCCTAACCTCTGTGAACAAGACAGCTT 1589
DB 3126 AATGAAATCCATGCTGCTCAGAGAGAGCTCCCTAACCTCTGTGTGAACAAGACAGCTT 3185
OY 1590 TGAGGAGCTCAGACCTGCGCTCTGCTGACACCCCGCTTCTGCTCCCTGCTGCTG 1649
DB 3186 AGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3231
OY 1650 CCGCATATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1709
DB 3232 TACTTGTCTTCTGTTGATTCCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 3290
OY 1710 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764
DB 3291 ATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3350
OY 1765 CCGTCTCTCAACCACTGATCTCT 1790
DB 3351 CCGTCTCTCAACCACTGATCTCT 3376

```

```

RESULT 13
AAS21297
ID AAS21297 standard; cDNA; 2343 BP.
AC AAS21297;
XX
XX 24-OCT-2001 (first entry)
XX

```

```

DE Human cDNA sequence encoding for PRO4404 polypeptide.
XX
XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
XX adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS Homo sapiens.
XX
XX WO200140466-A2.
XX
PD 07-JUN-2001.
XX
XX 01-DEC-2000; 2000MO-US32678.
XX
XX 01-DEC-1999; 99MO-US28301.
XX 01-DEC-1999; 99MO-US28634.
XX 02-DEC-1999; 99MO-US28551.
XX 02-DEC-1999; 99MO-US28564.
XX 02-DEC-1999; 99MO-US28565.
XX 09-DEC-1999; 99MO-US170262.
XX 16-DEC-1999; 99MO-US30095.
XX 20-DEC-1999; 99MO-US30911.
XX 20-DEC-1999; 99MO-US30999.
XX 30-DEC-1999; 99MO-US31243.
XX 06-JAN-2000; 2000MO-US00277.
XX 06-JAN-2000; 2000MO-US00376.
XX 11-FEB-2000; 2000MO-US03565.
XX 18-FEB-2000; 2000MO-US04341.
XX 18-FEB-2000; 2000MO-US04342.
XX 22-FEB-2000; 2000MO-US04414.
XX 24-FEB-2000; 2000MO-US04914.
XX 24-FEB-2000; 2000MO-US05004.
XX 01-MAR-2000; 2000MO-US05601.
XX 20-MAR-2000; 2000MO-US07377.
XX 21-MAR-2000; 2000MO-US07532.
XX 30-MAR-2000; 2000MO-US08439.
XX 17-MAY-2000; 2000MO-US13705.
XX 22-MAY-2000; 2000MO-US14042.
XX 30-MAY-2000; 2000MO-US14941.
XX 02-JUN-2000; 2000MO-US15264.
XX 10-NOV-2000; 2000MO-US30873.
XX
XX (GENTH ) GENENTECH INC.
XX
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
PI Gerritsen ME, Goddard A, Godowski RJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TR, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2001-408281/43.
XX
XX P-PSDB; AAU12225.
XX
XX Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical.
XX
XX Claim 3; Fig 107; 813pp; English.
XX
XX AAS21244-AAS21518 encode for novel human secretory and transmembrane
XX PRO polypeptides. The PRO polypeptides are useful to detect other
XX PRO polypeptides, to link bioactive molecules to cells expressing
XX PRO polypeptides, to modulate biological activities of cells expressing
XX PRO polypeptides, and to detect the presence of mammalian lung, colon,
XX breast, prostate, rectal, cervical or liver tumours by comparing PRO
XX polypeptide expression in a cell sample to that in a control sample.
XX Some of the 275 sequences are also useful to stimulate the release of
XX tumour necrosis factor-alpha (TNF-alpha) from human blood, the
XX proliferation or differentiation of chondrocytes, the proliferation or
XX gene expression in pericyte cells, the release of proteoglycans from
XX cartilage, the proliferation of inner ear utricular supporting cells or
XX of T-lymphocytes, the release of a cytokine from peripheral blood
XX monocytes (PBMCs), or the proliferation of endothelial cells. Some of

```

CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.

XX Sequence 2343 BP; 607 A; 581 C; 525 G; 630 T; 0 other;

Query Match 20.7%; Score 534; DB 22; Length 2343;

Best Local Similarity 61.5%; Pred. No. 1e-112; Matches 912; Conservative 0; Mismatches 560; Indels 12; Gaps 3;

```

QY 95 GGGCTCCCTGCTCATTCCTGCTGCTGATCAAGGCGATTCAGCTTACTGACAG 154
DB 281 GGGCTCCCTGCTCATTCCTGCTGCTGATCAAGGCGATTCAGCTTACTGACAG 154
QY 155 GGAGTGGCTGCTCAAGCCCTCGAGAGTCCCGTCCCTCCCTCCAGCTGCTTCCG 214
DB 341 GGAGTGGCTGCTCAAGCCCTCGAGAGTCCCGTCCCTCCCTCCAGCTGCTTCCG 400
QY 215 GCAATCCAGAGAGCTCAAGAGAGAGAGAGATCAAGAGATTCAGAAATGGGTGAGAG 274
DB 401 GCAATCCAGAGAGAGCTCAAGAGAGAGAGAGATCAAGAGATTCAGAAATGGGTGAGAG 457
QY 275 ATTCCCAAGTGCCTGCTCATTCCTGCTGATGGAGGCAAGTTCGTCAGCTTATGA 334
DB 458 ATTCCCAAGTGCCTGCTCATTCCTGCTGATGGAGGCAAGTTCGTCAGCTTATGA 517
QY 335 CCTTCACTATGAGAGATTCCTGGAGAGATCAAGAGAGAGATTCAGAAATGGGTGAGAG 394
DB 518 CCTTCACTATGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 577
QY 395 ATTCCGCTGCTCAAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 454
DB 578 ATTCCGCTGCTCAAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 637
QY 455 GCATGAGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 514
DB 638 GCATGAGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 697
QY 515 CATTGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 571
DB 698 CATTGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 757
QY 572 TTCCCTGCTGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 631
DB 758 TTCCCTGCTGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 817
QY 632 TGCCTTCAAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 691
DB 818 TGCCTTCAAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 877
QY 662 CATTGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 721
DB 878 CATTGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 937
QY 752 CAGCATCTAAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 811
DB 938 CAGCATCTAAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 997
QY 812 TGAGCAGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 871
DB 998 TGAGCAGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 1057
QY 872 GGAGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 931
DB 1058 GGAGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 1117
QY 932 GAGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 991
DB 1118 GAGAGAGATTCCTGGAGAGATTCAGAGAGAGATTCAGAGAGAGATTCAGAAATGGGTGAGAG 1177

```

```

QY 992 TGAAGCTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1051
DB 1178 TGAAGCTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1237
QY 1052 CCCAAGAGATTCAGAGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1111
DB 1238 CCCAAGAGATTCAGAGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1297
QY 1112 CATCACTGAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1171
DB 1298 CATCACTGAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1357
QY 1172 GAGGCTTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1231
DB 1358 GAGGCTTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1417
QY 1232 TGAAGCTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1291
DB 1418 TGAAGCTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1477
QY 1292 CAACCCAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1351
DB 1478 CAACCCAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1537
QY 1352 TGGCTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1405
DB 1538 TGGCTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1597
QY 1406 TGGGAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1465
DB 1598 TGGGAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1657
QY 1466 TGAAGCTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1525
DB 1658 TGAAGCTCAAGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1717
QY 1526 CAAAATGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1569
DB 1718 CAAAATGAGATTCCTGGAGAGATTCAGAGAGATTCAGAGAGATTCAGAAATGGGTGAGAG 1761

```

RESULT 14

ACA03656
 ID ACA03656 standard; cDNA; 2343 BP.

ACA03656;

AC ACA03656;

DT 23-MAY-2003 (first entry)

XX cDNA encoding human PRO polypeptide #54.

DE Human; PRO polypeptide; secreted and transmembrane protein;

KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;

KW differentiation; chondrocyte; tumour; genetic disorder;

KW cytosolic; gene; ss.

OS Homo sapiens.

PN US2003036180-A1.

XX US2003036180-A1.

XX 20-FEB-2003.

PD 09-MAY-2002; 2002US-0143114.

XX 31-MAR-1997; 97MO-US05230.

PR 12-JUN-1998; 98MO-US12456.

PR 14-JUL-1998; 98MO-US14552.

PR 28-AUG-1998; 98MO-US17888.

PR 10-SEP-1998; 98MO-US18824.

PR 14-SEP-1998; 98MO-US19093.

PR 14-SEP-1998; 98MO-US19094.

PR 14-SEP-1998; 98MO-US19177.

PR 14-SEP-1998; 98MO-US19330.

Db 458 ATACCTGCTGCTTCCCTTTCGATGAGGCGCTTACGACATTTTCGTATCATGA 517
Qy 335 CCTGATATATAGAGATATCTGGAGATCAGACCCGAAATCCAGATTCCTACAG 394
Db 518 CCCAGATATAGACACTTCTGAGCAGAAACAGATCCCAAGTCCAGATCTGAGAA 577
Qy 395 ATTCGCTGCTGATGATGAGGCTTGTCTCTCTGTGAATGGGACAGATGTTTCCA 454
Db 578 ATTCGACCTTCCACTTCTGAGAAAGACATGAGGCTCAGAGGACCCAGAGTTTCCA 637
Qy 455 GCATGACGAGATGCTGACCCGACCTTCCACTATGACATCTGAGGCCATGAGGGGCT 514
Db 638 GCATGCTGCTGCTACTACTCTCTGATTCATTTTAACTCTGAAAGCATACATGAGT 697
Qy 515 CAYGCAGACTCTGTACAGATGATGCTGACAAATGGGAAGAGCTCTTGGC---CAGA 571
Db 698 GATGCTGCTATCTGTAAATATGATGCTGATAGTGGAGAGATTTGACGACCTCAGA 757
Qy 572 TTCCCTCTGAGAGTTTTCAGACGCTCTCTGATGACCTTGGACACCATCATGAGTG 631
Db 758 CACAGCCGAGAGCTATAGACATCACTCATGATGCTCTGATATATCATGAAATG 817
Qy 632 TGCCTTACGCCATCAGGGGACATCCAGGTGACAGGAATCTCACTCTACATACAGC 691
Db 818 CGCTTTCAGCAGAGAACCACTGCGACAAACACACCCTCATGATCTTATGCAAAAGC 877
Qy 692 CATTAATGACTGAAACACTGGTTTTTCCGCTGAGGAATGCTTTTACACAGATGA 751
Db 878 CATATTGAACTCAGBAATCATATTTTCAACGCTGTACAGTTGTTGATACAGTGA 937
Qy 752 CACCATCTACAGCTGACCTCTGCTGCGCTGAGACACCGGCGCTGCGACGTGCGCA 811
Db 938 CATATTTTCAACTCAGCCCTCAGGGGCTACCGCTTCCAGAGTAAAGCGAGATGTTGA 997
Qy 812 TCAGACACACAGCAGATGATCCACTGAGAAAGCTCAACTACAGAGAGGGGAGCT 871
Db 998 TCAGTACACAGATACATATATCCAGAAAGAAAGAAATCCCTCCAGGCTGGGGTAAAGCA 1057
Qy 872 GGAGAGATCAAGAGAGAGGATTTGATTTTCTGATATCTCTCTGCGCAAAAT 931
Db 1058 GGATTAACACTCCGAGAGAGAGATCCAGAGATTTTCTGATATTTCTTCTTCTGCAAGGA 1117
Qy 932 GGAGATGAGACATCTGTGACAGAGACCTCCGCTGAGGTGAGACACGTTCACTGT 991
Db 1118 TGAAGTGGTACAGCTTCTCAGATATGATGACCTGGAAGTACAGACATCTCTGTT 1177
Qy 992 TGAGGCGCACAGACCAAGCAGAGTGGATCTCTGATCTCTATGCTTGCCACCA 1051
Db 1178 GGCAGAGACATGACACTTGGCAGAGCATCTCTGATCTTATCTGCTGCTGAA 1237
Qy 1052 CCCAGACATCAGAGAGGTGCGGAGAGATCCACACCTCTGAGGTGATGAGACCTC 1111
Db 1238 CCTGAGCATCAAGAGAGTCCGGAGAGAGGTCAGGGGATCTCTGGGGATGAGGCTTTC 1297
Qy 1112 CATCACTGGAACCACTGACACAGATGCTTACACACATGATGATTAAGAGGACAT 1171
Db 1298 TATCACTTGGACACAGCTGGGAGATGCTGACACCAATGTCATCAAGAGACGTG 1357
Qy 1172 GAGGCTTACCCAGCGGTGCCAGGCAATGGCAGAGAGCTCAGACATCTCCGTCACTTCC 1231
Db 1358 CCGATGATTTCTGAGAGTCCCTCCATTTTCCAGAGATCTCAGAACCCACTTACTTCCC 1417
Qy 1232 TGATGGGCGCTCTGCGCAAGATATGATGCTCTCTCATATTAAGGCTTACCA 1291
Db 1418 AGATGATGACATGCTGACGAGATCAACGGGTCTTATGATTTGGGGCTTACCA 1477
Qy 1292 CAACCCAAAATGTGGCCCAACCCAGAGGTGTTTGAACCTTTCGTTTGAACCGGGTTC 1351
Db 1478 CAACCTGTGTCTGAAAAAACCCAAAGGTCTTTGACCCCTTGAAGGTCTCTCAGAGAA 1537
Qy 1352 TGTCAACAGACCCAC-----GCTTCTGCTTCTTCTGAGAGATCAAGAACTGCAAT 1405

Db 1538 TTCTGATCAGAGACACCCCTATGCTACTTACATTTCTAGCTGATCAAGAACTGCAAT 1597
Qy 1406 TGGGAACAATTTGGCAATGAAAGAGCTGAAGTGGCCACGGCCCTGACCTGCTCGCTT 1465
Db 1598 TGGGACAGAGTTTGGCATGATGATTAAGTAAAGTAAACCATTTGCTGCTCCACTT 1657
Qy 1466 TGAGCTGTGCTGCTGATCCCAACAGAGATCCCATCCCATTTGACAGACTGTGTGAATC 1525
Db 1658 CAGAGTACTTCCAGACCCCAACAGGCTCTTACTTTCCCAACCATTTTACTCAAGCC 1717
Qy 1526 CAAAATGGAATCCACTGCTCTCAGAGAGCTCCCTAACCTT 1569
Db 1718 CAAGATGGAGTATTTTGCACTGAAAGAACTCTGAAATGTT 1761

RESULT 15
ACA04077
ID ACA04077 standard; cDNA, 2343 BP.
XX ACA04077;
AC
XX
XX
DT 27-MAY-2003 (first entry)
XX
DE Human cDNA encoding a secreted/transmembrane protein, SRQ ID 107.
XX
XX Human; ss; gene; secreted protein; transmembrane protein; PRO;
XX inflammatory disease; organ failure; atherosclerosis; cardiac injury;
XX infertility; birth defects; premature aging; AIDS; biosensor;
XX acquired immunodeficiency syndrome; cancer; diabetic complication;
XX bioreactor; tumor.
XX
OS Homo sapiens.
XX
XX
PN US2003032155-A1.
XX
PD 13-FEB-2003.
XX
XX
PF 03-MAY-2002; 2002US-0137865.
XX
XX
PR 31-MAR-1997; 97WO-US05230.
PR 12-JUN-1998; 98WO-US12456.
PR 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25106.
PR 05-JAN-1999; 98WO-US25108.
PR 08-MAR-1999; 98WO-US05028.
PR 10-MAR-1999; 98WO-US05190.
PR 13-SEP-1999; 98WO-US20944.
PR 15-SEP-1999; 98WO-US21090.
PR 15-SEP-1999; 98WO-US21547.
PR 05-OCT-1999; 98WO-US23089.
PR 29-NOV-1999; 98WO-US28214.
PR 30-NOV-1999; 98WO-US28313.
PR 30-NOV-1999; 98WO-US28409.
PR 01-DEC-1999; 98WO-US28301.
PR 01-DEC-1999; 98WO-US28634.
PR 02-DEC-1999; 98WO-US28651.
PR 02-DEC-1999; 98WO-US28654.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: February 15, 2004, 01:20:39 ; Search time 5297 Seconds

(without alignments)
11819.583 Million cell updates/sec

Title: US-09-942-429A-6
Perfect score: 2576
Sequence: 1 gaattccgcgcagatccagc.....aaaaaaacgcgaattc 2576

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthm:*
3: em_estin:*
4: em_estm:*
5: em_estm:*
6: em_estm:*
7: em_estro:*
8: em_estro:*
9: gb_est1:*
10: gb_est1:*
11: gb_est1:*
12: gb_est1:*
13: gb_est1:*
14: gb_est1:*
15: em_estfun:*
16: em_estfun:*
17: em_estfun:*
18: em_estfun:*
19: em_estfun:*
20: em_estfun:*
21: em_estfun:*
22: em_estfun:*
23: em_estfun:*
24: em_estfun:*
25: em_estfun:*
26: em_estfun:*
27: em_estfun:*
28: gb_est1:*
29: gb_est1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1012.8	39.3	2084	11	AK002528	Mus muscu
2	974.4	37.8	1201	13	BX340590	BX340580
3	769.4	29.9	854	12	B1759294	B1759294
4	660.8	25.7	1008	13	BX452362	BX452362

Result No.	Score	Match	Length	DB	ID	Description
5	631.2	24.5	1201	13	BX422210	BX422210
6	626.4	24.3	635	12	B1759809	B1759809
7	626	24.3	701	13	BX097936	BX097936
8	592.6	23.0	781	10	BG533264	BG533264
9	540.2	21.0	3860	11	AK038526	AK038526
10	522.6	20.9	792	12	BG973046	BG973046
11	513.4	19.9	844	12	B1148819	B1148819
12	510.8	19.8	872	12	BF687317	BF687317
13	499.2	19.4	834	12	B1147726	B1147726
14	490.4	19.0	670	10	CB455420	CB455420
15	482.2	18.7	806	12	B1143913	B1143913
16	481.2	18.7	779	12	B1102811	B1102811
17	478.2	18.5	495	14	CB163562	CB163562
18	472	18.3	835	12	BG970447	BG970447
19	470.2	18.3	864	12	B1103261	B1103261
20	466	18.1	737	14	CB952673	CB952673
21	463.4	18.0	864	12	B1219067	B1219067
22	463	18.0	838	9	A1097751	A1097751
23	462.8	18.0	828	9	A1956431	A1956431
24	460.8	17.9	833	12	BG969653	BG969653
25	456.4	17.7	779	14	CB600740	CB600740
26	454.4	17.6	761	10	BF383974	BF383974
27	449.6	17.5	711	14	CB950195	CB950195
28	447	17.4	746	10	BF788061	BF788061
29	444.8	17.3	876	12	B1101160	B1101160
30	439.2	17.0	599	14	CB457963	CB457963
31	435.8	16.9	782	10	BF782104	BF782104
32	433.6	16.8	733	14	CB954270	CB954270
33	433.6	16.8	831	12	BG970019	BG970019
34	430.8	16.7	745	10	BF788511	BF788511
35	430.6	16.7	761	14	CB950710	CB950710
36	428.8	16.6	753	10	BF783236	BF783236
37	428.4	16.6	696	12	B1331167	B1331167
38	426.6	16.6	769	12	B1221979	B1221979
39	424.2	16.5	731	14	CB955255	CB955255
40	424.2	16.5	898	10	BF533291	BF533291
41	424	16.5	649	14	CB417720	CB417720
42	422	16.4	572	10	BF043014	BF043014
43	418.6	16.2	926	10	BF532415	BF532415
44	417.2	16.2	737	14	CB600339	CB600339
45	416.2	16.2	658	12	B1148789	B1148789

ALIGNMENTS

RESULT 1
AK002528
LOCUS
DEFINITION Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:0610011D16 product:cyclochrome P450, 4a10, full insert sequence.
ACCESSION AK002528
VERSION AK002528.1 GI:12832575
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL MEDLINE
MEDLINE 99279253
PUBMED 10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL MEDLINE
PUBMED 11042159

REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE PUBMED	REFERENCE AUTHORS
3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Katsunai, T., Teshiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazada, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakauechi, S., Ikegami, T., Kasahigaki, K., Fujitake, K., Inoue, K., Ozawa, Y., Tazawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Toge, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		Kawai, J., Shitagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.
	RIKEN Integrated Sequence analysis (RISA) system -384 format		
	Sequencing pipeline with 384 multicapillary sequencer		
	Genome Res. 10 (11), 1757-1771 (2000)		
	11076861		4

TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851

AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6. (bases 1 to 2084)
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,

TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'] GAGAGAGAGCGCGCCGCACTCGAGTTTTTTTTTTTTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase

FEATURES

Source	Location/Qualifiers
1.	2084

CDS

```

/ctsnse type="kidney"
/clone lib="Riken full-length enriched mouse cDNA library"
/dev stage="adult"
37. 1566
/note=unnamed protein product; cytochrome P450, 4a10
(MD|MGJ.88611, GB|AK002528, evidence: BLASTN, 100%,
match=2084)
putative"
/codon_start=1
/protein_id="BA822165.1"
/db_xref="GI:12832576"
/db_xref="MGJ:88611"
/translation="MSTVSALSPTRPADSLSGFQVAVSLGLILLVRAVCFYLRHOWM
LKKFOOPSPDPFPMFGHEKFGQDOELQELIVSCENPSPAPRFMSKALITVADP
YMLVILGRSDPKANGAYRIAPMIGVGLLILNGPQPOHRMLTPAFAYDILKVPYKA
MADSLILMLKMERLADODSIRIPOHISLMTLDTWAKCAFSHGASVQVNDNRYTL
AIDIANLMLFHSRPNRIPTHODDTYKLSNSRNLAKOAKLADHDHGVITLKRKDOLOD
GELKRIKKRRRLDPDLILPFRMENGDSKMDKDLRAVDPMFEGSHOTTSGVSRIRP
ALAHVHQQRCEBVSQSLADGSSITMDHLDLPITWCIKELRLRYPFGVIREI
STVSTPDPDGRSLPRGVQVLTSLYGLNHPVPMDEVPDPSRPADSRHSRSLPFRF
GARNCTGKQFAMSELKVIALTLTRELLLPDPTVRVMPILARVLKSKNGIYLHKDL
H"

```

BASE COUNT	538	547	445	554
ORIGIN	a	c	g	c

Query Match	39.3%	Score 1012.8	DB 11	Length 2084
Best Local Similarity	78.7%	Pred. No. 8.2e-99		
Matches 1222	Conservative	0	Mismatches 327	Indels 3
				Gaps 1

9 CAGAGATCCAGCAGGTGCTGCACCATGAGTGTCTCTGTGCTGAGCCCCAGCAGACTTCTTG 68

Db 13 CAGCAACAGAGAGTGGATGACTATGAGTGTCTCTGCTCTAAGCCCAACCGATTCG / 2

[illegible]

22

[illegible]

240

1 0 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040

249 CACCGGATTTCAGGAAATCGGAGGAGCATTTCCCAAGTGCCTGTCCTCATYGGCTATGGGGA 308

Db 250 CAGGAATTTGTCTGTCATAGAGAACTTCCCAAGTGCTTCCCTAGATGTTCTGGGA 309

309 GGCAAGTTGTCGTCACGCTCTATGACCCCTGACTATATGAGGTGATTCGGGAGATCA 368

Db 310 AGCAAGCCTACTTAACTGTCTATGACCCCTGACTACATGAAGTGATTCTGGGGCGATCA 369

369 GACCCGAATCCCATGTTCTTACAGATTCTGGCTCCATGGATTGGGTACGGCTTGTCTC 428

Db 370 GATCCAAGGCTAATGCTGCTACAGATTGCTAGCTCCCTTGGATTGGATATGGTTTCTC 429

Oy	429	CTGTGAATGGGCACATGGTTCAGAGATGAGAGAGTGTGAACCCAGCCTTCCACAT	488
Db	430	TGTGTGAATGACACGCATGGTTCAGACAAGAGAGATGTACCCAGCCTTCCACAT	489
Oy	489	GACATCTGGAAGCCCTATGTGGGGCTCATAGGAGACTGTGTACAGATGTGTGACAA	548
Db	490	GACATCTGGAACCTATGTGAAAAACATGGCCGACTCATTCGATTGATGTGACAAA	549
Oy	549	TGGGAAGAGCTCTTGGCCAGGATTCCTCTGTGAGAGTCTTTCAGACAAGTCTCTTGATG	608
Db	550	TGGGAACGGCTGGCTGATCAGAGACTCTCTATAGAGATCTTTCACAACATCTCTTAATG	609
Oy	609	AACCCGGAACACATATGATGATGTGCTTCAGCCCATCAGAGAGACATCCAGGTGACAG	668
Db	610	AACCTAGACATGTATGATAATGTGCTTCAGCCCAAGAGGAGAGTTCAGGTATATGGA	669
Oy	669	AATTCAGTCCCTACATACAGGCACTTAATGATGACTGAAACAACCTGGTTTTTCCCGTGG	728
Db	670	AATTACAGAGACTTACTCCAGGCCATTGGGACTGTAATTAACCTATTTCATCTCCGTGTG	729
Oy	729	AGGAATGCTTTTCACCAAGATGACACATCTACAGCTGACCTCTGCTGGCCGTGACA	788
Db	730	AGGAATATCTTTCATCAGATGATACATCTTCAAACTTCTTCATATGGCCGCTTGACC	789
Oy	789	CACCGCGCTGCACAGTGGCCCATCAGACACAGACCAAGTATCCAACTGAGAAAGCT	848
Db	790	AAACAGAGTGTTCACATCTGCCCATATCACACAGATGAGTATCACTGAGAAAGAT	849
Oy	849	CAACTACAGAAAGAGGGGAGCTGAGAGATCAAGAGGAGAGGCACTTGGATTTCTG	908
Db	850	CAGCTGCAGATGAGGAGAGCTGGAAAGATCAAGAGAGAGAGCGTTTGATTTCTG	909
Oy	909	GATATCTCTCTTGGCCAAATGAGAGATGGAGACATCTGTCAACAAGACCTCCGT	968
Db	910	GATATCTCTTATTGTCAGAAATGAGAGATGGGACACAGATGTCTGACAGAGACTAGT	969
Oy	969	GCTGAGGTGACACCTTATGTTTGAAGGCCACAGACACAGCCAGTGGGATCTCTGG	1028
Db	970	GCTGAGGTGACACATATGTTTCAGAGGCCATGACACACAGCCAGTGGAGTCTCTGG	1029
Oy	1029	ATCCCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCGGAGAGATCCAC	1088
Db	1030	ATCTTCTATGCTCTGGCCACACCCCTGATTCACACAGAGATGAGAGAGAGTTGAG	1089
Oy	1089	AGCCTCTCGGGTGAATGAGACCTTCATCACTGGAACCACTGGACCAAGTGGCTTACCC	1148
Db	1090	AGCCTCTCGGGGAGATGATCTCTATACCTGTGATCACTAGACCAATTCCTATAC	1149
Oy	1149	ACCATGTGCATTTAAGAGAGCACTGAGGCTTACCAACCGGTGCCAGGCAATGGCAGAG	1208
Db	1150	ACCATGTGCATTCAGAGAGGCCCTTAAGGCTTACCAACCTGTGCCAGGCAATGTCCAGAGA	1209
Oy	1209	CTCAGCACTCCCGTACCTTCCCTGATGAGGGGGGCTCTGGCCCAAGGATCATGTGCTCT	1268
Db	1210	CTCAGTACATCTGTACCTTCCCTGATGAGCGCTTTAACCAAGGATGTCCAGGTCAACA	1269
Oy	1269	CTCTTCATTTATGACCTTCAACAACAACCAAAAGTGTGGCCAAACCAAGAGTGTTTGAC	1328
Db	1270	CTCTTCATTTATGATGTCCACAACAACCGAAGGTGTGGCCAAATCCAGAGTGTTTGAC	1329
Oy	1329	CCTTTCCGTTTGGACCGGGTGTGTCTCAACAACGACACGCTTCTCTGCTCTTCAGGA	1388
Db	1330	CCTTTCCAGTTTGGACACAGACTCTCCCGACACAGCACTCAATTCCTGCCCTTTCAGGA	1389
Oy	1389	GGATCAAGAACTGCAATGGGAAACAATTTGCAAGAAAGAGCTGAAGTGTCCAGGCC	1448
Db	1390	GGAGCAAGAACTGCAATGGGAAACAATTTGTATGATGAGCTGAAGTGTGTGACC	1449
Oy	1449	CTGAACCTGTCTCCGCTTTGAGCTGTGCTGTATCCCAACAGGATCCCATCCCATTTGCA	1508
Db	1450	CTGAACCTGTCTCCGCTTTGAGCTGTGCAATCTGCAAGATCTCTCAAGGATCCCATGCTTACCC	1509
Oy	1509	CGACTGTGTGAATCCAAAAATGAATTCACCTGCGTCTTCAGAGGCTCC	1560

DB	1510	CGACTTGTGTAAGTCCAAAAATGGAGATCTACTACATCTCAAGAGCTCC	1561
RESULT 2			
LOCUS	BX340590	1201 bp	mRNA linear EST 02-MAY-2003
DEFINITION	BX340590 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED		
ACCESSION	BX340590	Homo sapiens cDNA clone CS0DJ003Y106 5-PRIME, mRNA sequence.	
VERSION	BX340590.1	GI:30335923	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 947.r For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DJ003Y106&cluster=947.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DJ003BE030P1.		
FEATURES			
Source	1. 1201 /organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /clone="CS0DJ003Y106" /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED" /cell_line="JURKAT" /clone_1b="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT.6 vector. Library was normalized."		
BASE COUNT	261 a 311 c 326 g 257 t 46 others		
ORIGIN			
Query Match	37.8%; Score 974.4; DB 13; Length 1201;		
Best Local Similarity	94.0%; Pred. No. 1.3e-94;		
Matches 998; Conservative	21; Mismatches 42; Indels 1; Gaps 1;		
Db	7 CGCAGAGATCCAGCAGGTGCTGCACCAATGATGTCCTTGTGCTGAGAGCCGCCAGCAGATCC 66		
Db	58 CTGAGAGATCCAGCAGGKCTGCACCAATGATGTCCTTGTGCTGAGAGCCGCCAGCAGATCC 117		
Db	67 TGGGTATATGTCCTGGAATCCTCCAAAGCGAGCTCCCTGCTCATTCCTGCTTGTGCTGCTGA 126		
Db	118 TGGGTATATGTCCTGGAATCCTCCAAAGCGAGCTCCCTGCTCATTCCTGCTTGTGCTGCTGA 177		
Db	127 TCAAGGAGTTGACCTTCACTTCAGCAGGAGGTGCTCAAGCCCTCCAGCAGTTCC 186		
Db	178 TCAAGGAGTTGACCTTCACTTCAGCAGGAGGTGCTCAAGCCCTCCAGCAGTTCC 237		
Db	187 CGTCCCTCCCTCCCACTGCTCTTGTGGGAGCATCCAGAGGCTTCCACAGAGCCAGAGAGC 246		
Db	238 CGTCCCTCCCTCCCACTGCTCTTGTGGGAGCATCCAGAGGCTTCCACAGAGCCAGAGAGC 297		
Db	247 TACAACGAGTTCCAGAAATGGATGGAGAGCATTCACCAAGTGCCTGTCCTATTTGGCTATGGG 306		
Db	298 TACAACGAGTTCCAGAAATGGATGGAGAGCATTCACCAAGTGCCTGTCCTATTTGGCTATGGG 357		

QY 307 GAGGCAAGTGTGTCAGCTCTATGACCTGATATATGAAAGTGTCTTGGGAGAT 366
 DB 358 GAGGCAAGTGTGTCAGCTCTATGACCTGATATATGAAAGTGTCTTGGGAGAT 417
 QY 367 GAGGCAAGTGTGTCAGCTCTATGACCTGATATATGAAAGTGTCTTGGGAGAT 426
 DB 418 GAGGCAAGTGTGTCAGCTCTATGACCTGATATATGAAAGTGTCTTGGGAGAT 477
 QY 427 TCCTGTGAATGGGAGAGATGTTCCAGATGAGAGATGAGAGATGAGAGAT 486
 DB 478 TCCTGTGAATGGGAGAGATGTTCCAGATGAGAGATGAGAGATGAGAGAT 537
 QY 487 ATGACATCTGAAGCCCTATGAGAGATGAGAGATGAGAGATGAGAGAT 546
 DB 538 ATGACATCTGAAGCCCTATGAGAGATGAGAGATGAGAGATGAGAGAT 597
 QY 547 AATGGGAAGAGCTCTTGGGAGAGATGAGAGATGAGAGATGAGAGAT 606
 DB 598 AATGGGAAGAGCTCTTGGGAGAGATGAGAGATGAGAGATGAGAGAT 657
 QY 607 TGAACCTGAGACCATATGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 666
 DB 658 TGAACCTGAGACCATATGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 717
 QY 667 GGAATTTCTGAGTCTTATGAGAGATGAGAGATGAGAGATGAGAGAT 726
 DB 718 GGAATTTCTGAGTCTTATGAGAGATGAGAGATGAGAGATGAGAGAT 777
 QY 727 TGAAGAAATGCTTTCAGAGAGATGAGAGATGAGAGATGAGAGAT 786
 DB 778 TGAAGAAATGCTTTCAGAGAGATGAGAGATGAGAGATGAGAGAT 837
 QY 787 CACACCGGCTGCTGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 846
 DB 838 CACACCGGCTGCTGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 897
 QY 847 CTCACATGAGAGAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 906
 DB 898 CTCACATGAGAGAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 957
 QY 907 TGAATATCTCTCTTTCAGAGAGATGAGAGATGAGAGATGAGAGAT 966
 DB 958 TGAATATCTCTCTTTCAGAGAGATGAGAGATGAGAGATGAGAGAT 1017
 QY 967 GAGCTGAGAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1026
 DB 1018 GAGCTGAGAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1076
 QY 1027 GAGCTGAGAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1068
 DB 1077 TCTATATCTCTCTTTCAGAGAGATGAGAGATGAGAGATGAGAGAT 1118

RESULT 3
 BI759294 854 bp mRNA linear EST 25-SBP-2001
 LOCUS 603042936F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183289 5',
 DEFINITION mRNA sequence.
 ACCESSION BI759294
 VERSION BI759294.1 GI:15750872
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 854)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: gsbds-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.

FEATURES
 source
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.jnl.gov>
 Plate: L1M11457 row: h column: 10
 High quality sequence stop: 830.
 Location/Qualifiers
 1..854
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5183289"
 /lab_host="DH10B"
 /lab_lib="NIH_MGC_116"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: ScaRV (destroyed); RNA
 source anonymous pool of 3 clones, age 26 yo male, 49 yo
 female, 71 yo male, 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH-MGC Library."

BASE COUNT 187 a 252 c 223 g 192 t
 ORIGIN

Query Match 29.9%; Score 769.4; DB 12; Length 854;
 Best Local Similarity 98.5%; Pred. No. 8.2e-73; Indels 7; Gaps 6;
 Matches 840; Conservative 0; Mismatches 6;

QY 7 CGCAGAGATCAGACAGAGTGTGACACATGAGTCTCTGTGAGAGAGAGAGAGAT 66
 DB 6 CTCAGAGATCAGACAGAGTGTGACACATGAGTCTCTGTGAGAGAGAGAGAT 65
 QY 67 TGGGTATGTCCTGGAATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAT 126
 DB 66 TGGGTATGTCCTGGAATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAT 125
 QY 127 TCAAGCAGTTGAGCTTACCTGACAGAGAGAGAGAGAGAGAGAGAGAT 186
 DB 126 TCAAGCAGTTGAGCTTACCTGACAGAGAGAGAGAGAGAGAGAGAGAT 185
 QY 187 CGTGGCTCTCCCTCCAGCTGAGCTCTTGGGAGATCAGAGAGAGAGAGAT 246
 DB 186 CGTGGCTCTCCCTCCAGCTGAGCTCTTGGGAGATCAGAGAGAGAGAGAT 245
 QY 247 TACAAGGATTCAGAAATGGGTGAGACATTCAGAGAGAGAGAGAGAT 306
 DB 246 TACAAGGATTCAGAAATGGGTGAGACATTCAGAGAGAGAGAGAGAT 305
 QY 307 GAGGCAAGTGTGTCAGCTCTATGACCTGATATATGAAAGTGTCTTGGGAGAT 366
 DB 306 GAGGCAAGTGTGTCAGCTCTATGACCTGATATATGAAAGTGTCTTGGGAGAT 365
 QY 367 CAGACCGGAATCCAGAGTTCCTACAGATTCAGAGATTCAGAGATTCAGAGAT 426
 DB 366 CAGACCGGAATCCAGAGTTCCTACAGATTCAGAGATTCAGAGATTCAGAGAT 425
 QY 427 TCCTGTGAATGGGAGAGATGTTCCAGATGAGAGATGAGAGATGAGAGAT 486
 DB 426 TCCTGTGAATGGGAGAGATGTTCCAGATGAGAGATGAGAGATGAGAGAT 485
 QY 487 ATGACATCTGAAGCCCTATGAGAGATGAGAGATGAGAGATGAGAGAT 546
 DB 486 ATGACATCTGAAGCCCTATGAGAGATGAGAGATGAGAGATGAGAGAT 545
 QY 547 AATGGGAAGAGCTCTTGGGAGAGATGAGAGATGAGAGATGAGAGAT 606
 DB 546 AATGGGAAGAGCTCTTGGGAGAGATGAGAGATGAGAGATGAGAGAT 605

QY 607 TGACCTGAGACATCATGAGTGGCTTCAGCATGAGGAGCATCAAGTGA-C 665
DB 606 TGACCTGAGACATCATGAGTGGCTTCAGCATGAGGAGCATCAAGTGA-C 665
QY 666 AGGATTCCTGAGTGGCTTCAGCATGAGGAGCATCAAGTGA-C 725
DB 666 AGGATTCCTGAGTGGCTTCAGCATGAGGAGCATCAAGTGA-C 725
QY 726 GTGAGGAATGCC-TTTCACCAAGATGACATCTACAGCTTCAGCTTCGCGCGCTG 784
DB 724 GTGAGGAATGCC-TTTCACCAAGATGACATCTACAGCTTCAGCTTCGCGCGCTG 782
QY 785 GACACACCGCGCTTCAGCTTCAGCATGAGGAGCATCAAGTGA-C 844
DB 783 GACACACCGCGCTTCAGCTTCAGCATGAGGAGCATCAAGTGA-C 840
QY 845 GGCTCACTACAG 857
DB 841 GGCTCACTACAG 853

RESULT 4
BX452362 1008 bp mRNA linear EST 22-MAY-2003
LOCUS BX452362 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM002YC01 5-PRIME, mRNA sequence.
ACCESSION BX452362
VERSION BX452362.1 GI:31024367
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1008)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 947.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG0632B03_CS05970_1&cluster=947.r.
Contact: Peng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Parade Avenue Genoscope sequence ID: CS0BAG0632B03_CS05970_1.
Location/Qualifiers
1. 1008
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM002YC01"
/cissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_1lb="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

BASE COUNT 203 a 311 c 226 g 267 t 1 others
ORIGIN

Query Match 25.7%; Score 660.8; DB 13; Length 1008;
Best Local Similarity 87.7%; Pred. No. 2.3e-61;
Matches 780; Conservative 0; Mismatches 98; Indels 11; Gaps 5;

QY 1118 CTGGAACCACTGAGACAGATGCCCTTACACACCATGTCATTAAAGAGGCACTGAGGCT 1177
DB 49 CAGGAACCACTGAGACAGATGCCCTTACACACCATGTCATTAAAGAGGCACTGAGGCT 108

QY 1178 CTACCCACCGGTGCCAGGCAATTGGCAGAGAGCTACGACTCCCTCCATCCCTTCATG 1237
DB 109 CTACCCACCGGTGCCAGGCAATTGGCAGAGAGCTACGACTCCCTCCATCCCTTCATG 168
QY 1238 GGGCTCTGTCCTGACCAAGATGATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1297
DB 169 GGGCTCTGTCCTGACCAAGATGATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 228
QY 1298 AAAAGTGTGGCCCAACCCAGAGGTGTTGAACCTTTCGGTTTGAACCGGGTTCGCTCA 1357
DB 229 AAAAGTGTGGCCCAACCCAGAGGTGTTGAACCTTTCGGTTTGAACCGGGTTCGCTCA 288
QY 1358 ACAAGCCAGCGCTTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1417
DB 289 ACAAGCCAGCGCTTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 348
QY 1418 TGGCATTGAACGAGCTGAAGGTGGCCAGCGGCTGACCTGCTCCGCTTGGAGCTGCTGCC 1477
DB 349 TGGCATTGAACGAGCTGAAGGTGGCCAGCGGCTGACCTGCTCCGCTTGGAGCTGCTGCC 408
QY 1478 TGATCCACACAGATCCCATCCCATTCGACGACTTGTGTAATCCAAAATGGAAT 1537
DB 409 TGATCCACACAGATCCCATCCCATTCGACGACTTGTGTAATCCAAAATGGAAT 468
QY 1538 CCACCTGCGTCTCAGAGGCTCCCTTACCCCTTGTGAAGACAGACAGCTTTGAGGCGC 1597
DB 469 CCACCTGCGTCTCAGAGGCTCCCTTACCCCTTGTGAAGACAGACAGCTTTGAGGCGC 528
QY 1598 TCCACCTGCGTCTCAGAGGCTCCCTTACCCCTTGTGAAGACAGACAGCTTTGAGGCGC 1657
DB 529 TCCACCTGCGTCTCAGAGGCTCCCTTACCCCTTGTGAAGACAGACAGCTTTGAGGCGC 587
QY 1658 CTGTTTCTGTCGACCACTTCCCTTCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCT 1717
DB 588 CTGTTTCTGTCGACCACTTCCCTTCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 647
QY 1718 GCGCTTCT 1775
DB 648 GCGCTTCT 707
QY 1776 CCCACTGTATCTCTGT-----TGGAGAAAAGCTGATGTTGGAGAAC--TGAGGC 1828
DB 708 CCCACTGTATCTCTGT-----TGGAGAAAAGCTGATGTTGGAGAAC--TGAGGC 767
QY 1829 CGAGCTGATGCTGACAT 1888
DB 768 CGAGCTGATGCTGACAT 827
QY 1889 T-AAAACCTCTGTGGCCCTTGGAAACCAAGCTCTGTGCTGAAGGTGGAAGCTTACC 1947
DB 828 T-AAAACCTCTGTGGCCCTTGGAAACCAAGCTCTGTGCTGAAGGTGGAAGCTTACC 887
QY 1948 TGAAGCACTAT 1996
DB 888 TGAAGCACTAT 936

RESULT 5
BX422210 1201 bp mRNA linear EST 15-MAY-2003
LOCUS BX422210 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM002YC01 5-PRIME, mRNA sequence.
ACCESSION BX422210
VERSION BX422210.1 GI:30764162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL
COMMENT

Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: sequef@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 947.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DM002AB010P1cluster=947.r. Contact :
Peng Liang Email: liliang@life.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DM002AB010P1.

FEATURES
source

1.1201
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CS0DM002AB010P1"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/note="Organ: liver; Vector: PCWSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the PCWSPORT 6
vector. Library was not normalized."
Location/Qualifiers
275 a 306 c 307 g 269 t 44 others

BASE COUNT
ORIGIN

Query Match 24.5%; Score 631.2; DB 13; Length 1201;
Best Local Similarity 98.8%; Pred. No. 2.8e-58;
Matches 656; Conservative 1; Mismatches 4; Indels 3; Gaps 2;

7 CCAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT-- 64
70 CTCAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACTCA 129
65 CCGAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 123
130 CTGAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 189
124 TATCAAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 183
190 TATCAAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 249
184 TCCCGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 243
250 TCCCGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 309
244 AGCTAAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 303
310 AGCTAAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 369
304 GGGAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 363
370 GGGAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 429
364 GATCAAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 423
430 GATCAAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACT 489
424 TGTCTGTGATGAGGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAG 483
490 TGTCTGTGATGAGGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAG 549
484 ACTATGATCTCTGAGAGCTCTATGTGGGCTCTATGTGGGCTCTATGTGGGCT 543
550 ACTATGATCTCTGAGAGCTCTATGTGGGCTCTATGTGGGCTCTATGTGGGCT 609
544 ACAATGAGAGAGCTCTGAGAGCTCTATGTGGGCTCTATGTGGGCTCTATGTGG 603
610 ACAATGAGAGAGCTCTGAGAGCTCTATGTGGGCTCTATGTGGGCTCTATGTGG 669

QY 604 TATGATCCTTGACACCATCATGATGATGCTTACGCCATCAGGCGAGATCCAGCTGG 663
DB 670 TATGATCCTTGACACCATCATGATGATGCTTACGCCATCAGGCGAGATCCAGCTGG 729
QY 664 ACAG 667
DB 730 ACAG 733

RESULT 6

LOCUS B1759809 635 bp mRNA linear EST 25-SEP-2001
DEFINITION 603045693F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186015 5',
mRNA sequence.

ACCESSION B1759809
VERSION B1759809.1 GI:15751387
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 635)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
CONTACT: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov

Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M1464 row: 1 column: 24
High quality sequence step: 632.
Location/Qualifiers

FEATURES
source

1.635
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5186015"
/lab_host="DH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
PCW-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

BASE COUNT
ORIGIN

Query Match 24.3%; Score 626.4; DB 12; Length 635;
Best Local Similarity 99.8%; Pred. No. 1.3e-57;
Matches 627; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 CGAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACTCC 66
DB 7 CTGAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACTCC 66
QY 67 TGGGTATGCTCTGAGATCTCTCAAGCGGCTCTCTGCTCAATCTGCTGCTGTA 126
DB 67 TGGGTATGCTCTGAGATCTCTCAAGCGGCTCTCTGCTCAATCTGCTGCTGTA 126
QY 127 TCAAGAGATCCAGAGAGTGTGACCATAGTGTCTGTGCTGAGCCGACAGACTCC 186

DEFINITION 601860630r2 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4072062 5', mRNA sequence.

ACCESSION BG533264

VERSION BG533264.1 GI:13524804

KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

Home sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE NIH-MGC http://mhc.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cga@bbs-rcmail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLMT at: http://image.llnl.gov

Plate: L10C918 row: k column: 07

High quality sequence step: 766.

Location/Qualifiers

1. 781

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4072062"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC_76"

/note="Organ: liver; Vector: pDNR-LIB (Clontech); site: 1: SfiI (ggcgccgcggcc); site 2: SfiI (ggcgccatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGACC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGACGGCCGACATG-3' (where B = A, C, G, or T). Average insert size 1.85 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH MGC library."

BASE COUNT 168 a 227 c 204 g 182 t

ORIGIN

Query Match 23.0%; Score 592.6; DB 10; Length 781;

Best Local Similarity 96.0%; Pred. No. 4.5e-54;

Matches 629; Conservative 0; Mismatches 24; Indels 2; Gaps 2;

QY 7 CGCAGAGATCCAGAGTGTCTGTCAGATGCTCTGTGTCAGAGCCCGCAGAGATCC 66

DB 19 CTCAAGATCCAGAGTGTCTGTCAGATGCTCTGTGTCAGAGCCCGCAGAGATCC 78

QY 67 TGGGTGATGTCCTGGAATCTCTCAAGCGGCTCTCTGTCATTTGCTTGTCTGA 126

DB 79 TGGGTGATGTCCTGGAATCTCTCAAGCGGCTCTCTGTCATTTGCTTGTCTGA 138

QY 127 TCAAGGAGTTCAGCTCTACCTGACAGAGGAGTGTCTCAAGGCTCTCAAGGCTCTCC 186

DB 139 TCAAGGAGTTCAGCTCTACCTGACAGAGGAGTGTCTCAAGGCTCTCAAGGCTCTCC 198

QY 187 CGTGGCTCTCTCTCCACTGCTCTTCTGGGCACTTCAGAGAGCTCCCAAGAGAC 246

DB 199 CGTGGCTCTCTCTCCACTGCTCTTCTGGGCACTTCAGAGAGCTCCCAAGAGAC 258

QY 247 TACAAAGGATTCAGAAATGGGTGAGACATTTCCAAAGTGTCTCTCATTTGGCTATGG 306

DB 259 TACAAAGGATTCAGAAATGGGTGAGACATTTCCAAAGTGTCTCTCATTTGGCTATGG 318

QY 307 GAGGGAAGTGTCTGTCCAG-CTCATGACCTGATATGAGAGGATTTCTGGGAGGA 365

DB 319 GAGGGAAGTGTCTGTCCAGCTCATGACCTGATATGAGAGGATTTCTGGGAGGA 378

QY 366 TCAGACCCGGAATCCCATGTTCTTACAGATTCCTGCTCCATGATT-GGTTACGGCTT 424

DB 379 TCAGACCCGGAATCCCATGTTCTTACAGATTCCTGCTCCATGATTGTTGGGATCGCTT 438

QY 425 GCTCCGTTGAATGGGACAGATGGTTCAGAGTCAGAGATGTCAGATGTCAGATGTCAG 484

DB 439 GCTCCGTTGAATGGGACAGATGGTTCAGAGTCAGAGATGTCAGATGTCAGATGTCAG 498

QY 485 CTATGACATCTGTAAGCCCTATATGAGGAGCTCATGAGCACTGTATGAGATGCTGA 544

DB 499 CTATGACATCTGTAAGCCCTATATGAGGAGCTCATGAGCACTGTATGAGATGCTGA 558

QY 545 CAAATGGGAAGAGCTCTTGGCCAGAGATTCCTCTGAGAGTCTTTCAGAGATGCTCTT 604

DB 559 CAAATGGGAAGAGCTCTTGGCCAGAGATTCCTCTGAGAGTCTTTCAGAGATGCTCTT 618

QY 605 GATGACCTCGAGACACATCATGATGATGTCCTTACAGCCATGAGGAGCATTCAG 659

DB 619 GATGACCTCGAGACACATCATGATGATGTCCTTACAGCCATGAGGAGCATTCAG 673

RESULT 9

AK038526

LOCUS AK038526

DEFINITION Mus musculus adult male hypothalamus cDNA, RIKEN full-length enriched library, clone:A230025G20 product:similar to SIMILAR TO CYTOCHROME P450, 4A family [Mus musculus], full insert sequence.

ACCESSION AK038526

VERSION AK038526.1 GI:26332620

KEYWORDS HTC; CAP trapper

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Carninci P. and Hayashizaki Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE

AUTHORS Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M. and Hayashizaki Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20493774

PUBMED 11042159

REFERENCE

AUTHORS Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Kono H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikogami T., Kashiwagi K., Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai D., Okazaki Y., Muramatsu M., Inoue Y., Kira A. and Hayashizaki Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE

AUTHORS Kawai D., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamane K., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kouchiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L. M., Staudt F., Suzuki R., Tomita M., Wagner J., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Botfield D., Boujunga N., Carninci P., de Bonaudo M. F., Brownstein M. D., Bull C.,

Fletcher, C., Fujita, A., Gariboldi, M., Gasthincich, S., Hill, D., Hofmann, M., Hune, D.A., Kamtaya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzei, J., Mombers, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Witterker, C., Wilmink, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashitaki, Y.

TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)

MBDLIN 21085660
PUBMED 11217851

QY 983 GTTCATGTTTGAAGGCGCAGACACCAAGCCAGTGGGATCTCTGATCTCTATGCTCT 1042
 DB 906 CTTCATGTTGAGGAGAGACATGATGCTCTGACACTGACATCTCTGCTTACTGCT 965
 QY 1043 GAGCACAACCCCAAGCATCAGAGAGGAGCGGAGAGATCAAGCCTCTGAGGTGA 1102
 DB 966 GAGCTTAATCCCGACATCAGAGACATGCGGACAGATCAGAGATCAGAGATCCTGAGAGA 1025
 QY 1103 TGAAGCTTCATCAGCTGGAACCACTGGAACCAAGATGCGCTTACACCAACCATGATTA 1162
 DB 1026 TGGGCTTCATCAGCTGGAACCACTGGAACCAAGATGATTAACCAACATGATTA 1085
 QY 1163 GAGAGCATGAGGCTTACCAACCGGTCAGGCAATGAGAGATGAGATCACTCCCT 1222
 DB 1086 GAGAGCATGAGGCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145
 QY 1223 CACCTTCCCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1282
 DB 1146 TACCTTCCCAAGTGAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
 QY 1283 CTTTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1342
 DB 1206 TCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1265
 QY 1343 ACCGGGCTTCTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1396
 DB 1286 TAAAGGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1325
 QY 1397 GAACTGCAATGAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1456
 DB 1326 GAACTGCAATGAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1385
 QY 1457 GCTCCGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1516
 DB 1386 GCTCACTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1445
 QY 1517 GTTGAATTCACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1569
 DB 1446 CCTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1498

RESULT 10
 LOCUS BG973046 792 bp mRNA linear EST 12-JUN-2001
 DEFINITION 602841036P1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:4975160
 5' mRNA sequence.
 ACCESSION BG973046 GI:14360683
 VERSION BG973046.1 GI:14360683
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 792)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Straube, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Jeffrey B. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: L14M10966 row: h column: 09
 High quality sequence start: 10
 High quality sequence stop: 789.
 Location/Qualifiers
 1..792
 /organism="Mus musculus"

FEATURES
 Source

/mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4975160"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1ib="NCI CGAP Kid14"
 /note="Organ: Kidney; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 BASE COUNT 214 a 207 c 185 g 186 t
 ORIGIN

Query Match 20.3%; Score 522.6; DB 12; Length 792;
 Best Local Similarity 80.4%; Pred. No. 1,1e-46;
 Matches 612; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 531 CGAGTGAATGCTGGAACAAATGGAAGAGCTCTTGGCCAGAGATTCCTCTGAGGTCTT 590
 DB 32 CGATTGATGCTGACAAATGGAACGCGTGGCTGATCAGAGATCTCTTATGAGATCTT 91
 QY 591 CAGCAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 650
 DB 92 CAACATCTCTCTATATGATGATGATGATGATGATGATGATGATGATGATGATG 151
 QY 651 AGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 710
 DB 152 AGTGTTCAGTGAATGGAATTAACAGATCTCTCAGGCTGATGAGGATGATGATG 211
 QY 711 CTGATTTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 770
 DB 212 CTATTTCACTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 271
 QY 771 TCTGCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 830
 DB 272 TCAATGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 331
 QY 831 ATCCAACTGAGAGAGCTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 890
 DB 332 ATCAAGCTGAG 391
 QY 891 AGGATTTGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 950
 DB 392 AGAGCTTGAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 451
 QY 951 TCAGACAGAGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1010
 DB 452 TCGACAGAGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 511
 QY 1011 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1070
 DB 512 GCCAGTGAATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 571
 QY 1071 TGGCGGAGAGATTCACAGAGCTCTGATGATGATGATGATGATGATGATGATGATG 1130
 DB 572 TCGAGAGAGAGATTCAGAGCTCTGATGATGATGATGATGATGATGATGATGATG 631
 QY 1131 GACCAAGTCTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1190
 DB 632 GACCAAGTCTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 691
 QY 1191 CAGAGATTTGAG 1250
 DB 692 CAGAGATTTGAG 751
 QY 1251 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1291
 DB 752 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 792

RESULT 11
 LOCUS B1148819 844 bp mRNA linear EST 05-JUL-2001

DEFINITION 602912282F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5053028 5',
 RNA sequence.
 ACCESSION B1148819
 VERSION B1148819.1 GI:14608820
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 844)
 NIH-MGC http://mgi.mc.man.ac.uk/;
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rtm.man.ac.uk
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M11144 row: d column: 21
 High quality sequence stop: 815.
 Location/Qualifiers
 1..844
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5053028"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1ib="NCI CGAP L19"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 233 a 195 c 203 g 213 t
 ORIGIN
 Query Match 19.9%; Score 513.4; DB 12; Length 844;
 Best Local Similarity 77.9%; Pred. No. 1e-45;
 Matches 657; Conservative 0; Mismatches 181; Indels 5; Gaps 3;
 233 ACAGGACGAGAGCTACAAACGATTCAGAAATGGGTGAGAGATTCCTCAAGTCCGCTGTC
 Db 3 AAAGGACGAGAGCTTCAGAAATTCCTCAAGTTCAGAAATTCCTCAAGTCCGCTGTC 62
 293 TCATTGGCTATGGGAGGCAAGTTCGTGTCCAGCTCTATGACCTGACTATATGAAGT 352
 Db 63 ACAGTGGCTATGGGAGGCAAGTTCGTGTCCAGCTCTATGACCTGACTATATGAAGT 122
 353 GATTCGGGAGATGAGACCCGAATCCAGATGCTTCATGATTCCTGCTCCATGAT 412
 Db 123 GATTCGGGAGATGAGACCCGAATCCAGATGCTTCATGATTCCTGCTCCATGAT 182
 413 TGGGTACGGCTGCTCTGTGATGGGAGATGCTTCAGATTCAGATTCAGATTCAGAT 472
 Db 183 TGGGTACGGCTGCTCTGTGATGGGAGATGCTTCAGATTCAGATTCAGATTCAGAT 242
 473 CCCAGCTTCATATGATTCATGAGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAG 532
 Db 243 CCCAGCTTCATATGATTCATGAGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAG 302
 533 AGTATGCTGACAAATGAGGAGAGCTCTTGGCCAGATTCCTCTGAGAGTCTTTGA 592
 Db 303 TGTATGCTGACAAATGAGGAGAGCTCTTGGCCAGATTCCTCTGAGAGTCTTTGA 362
 593 GAGAGCTTCCTGATGAGCTTCGAGACCATCATGAGTGGCTTCAGCATCAGGAGGAG 652
 Db 363 ACACATTCACCTTATGAGCTTCGAGACCATCATGAGTGGCTTCAGCATCAGGAGGAG 422
 653 CATCATGATGACAGAAATTCAGTCTCTATCATACAGGCTTATGAGCTTATGAGCAACCT 712

Db 423 TGTCCAGTGGACAGAAATTCAGATTCATATCCAGGAGTGGAGCTTGAACATCT 482
 713 GGTATTTTCCCGTGTGAGGATATGCTTTTCACGAGATGACACCATCTACAGCTGACCTC 772
 Db 483 CGTTTTCCTGTCGAGGAGATCATCTTTTACCGAGATGACATCATCTTACAGATGCTC 512
 773 TGTGAGCTGAGACACCGGCTGCGAGCTGCGCATGACACACAGACCAAGTAT 832
 Db 543 TATGAGCTGAGACACCGGCTGCGAGCTGCGCATGACACACAGACCAAGTAT 602
 833 CCAACTGAGGAGGCTGAGACCTACAGAGGA-GGGGAGCTGAGAGATCAGAGGAGA 891
 Db 603 CAAATCAGAGGAGATTCATCTCAGATGACGAGAGTGGAAAGCTTAAACGACCA 662
 892 GGCATTTGATTTTCTGATATCTCTCTGCGCCAAATGAGAAATGAGAGCATCTGT 951
 Db 663 GGC-GATGATTTTCTGATATCTCTCTGCGCCAAATGAGAAATGAGAGCATCTGT 721
 952 CAGACAGAGACCTCGGTGAGTGGAGACGCTTCATGTTTGAAGGCGCACGACACAG 1011
 Db 722 CTGATAGAGACCTCGGTGAGTGGAGACGCTTCATGTTTGAAGGCGCACGACACAG 781
 1012 CCAGTGGATCT---CTGATCTCTATGCTCTGCGCCACACACCCAGCATCAGAGA 1068
 Db 782 GCTAGTGGAAATCTCTGAGTCTTCTATGCTTGGGCAAAATTCCTGAACTTCAGAGA 841
 1069 GGT 1071
 Db 842 GAT 844
 RESULT 12 872 bp mRNA linear EST 22-DEC-2000
 BF687317
 LOCUS 602102348F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4220771
 DEFINITION
 5', mRNA sequence.
 ACCESSION BF687317
 VERSION BF687317.1 GI:11972714
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 872)
 NIH-MGC http://mgi.mc.man.ac.uk/;
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rtm.man.ac.uk
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M9804 row: o column: 12
 High quality sequence stop: 733.
 Location/Qualifiers
 1..872
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4220771"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1ib="NCI CGAP Kid14"
 /note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 218 a 231 c 211 g 212 t

ORIGIN

Query Match 19.8%; Score 510.8; DB 10; Length 872;
 Best Local Similarity 80.2%; Pred. No. 1.9e-45;
 Matches 624; Conservative 0; Mismatches 152; Indels 2; Gaps 2;

575 CCTCTGAGAGTCTTTCAGACAGTCTCTTGAATGACCTTGAACCATCATGAAAGTTC 634
 1 CTCTTAGAGATCTTTCACACATCTCTTAATGACCTTGAACCATCATGAAAGTTC 60

635 CTTCAGCATCAGGGGAGGATCCAGGTGAGAGAAATTCAGTCTTACATACAGGCAAT 694
 61 CTTCAGCATCAGGGGAGGATCCAGGTGAGAGAAATTCAGTCTTACATACAGGCAAT 120

695 TAGTGAACCTGAAACAACCTGTTTTCCTGTTGAGAAATGCTTTCACAGAAATGAC 754
 121 TGGGGATCTTGAATTAACCTATTCACCTCCGTTGAGAAATTCATTCAGAAATGAC 180

755 CATCTACAGCTGACCTTCTGCTGAGGCTGAGACACCGGCTGACAGTGGCCATCA 814
 181 CATCTACAAATCTTCTTCAATGCGCTTGGCCAAACAGCTTGTCAACTTGGCCATCA 240

815 GACACAGACCAAGATCCAACTGAGAAAGCTCAACTACAGAGAGGGAGGCTGAG 874
 241 TCACACAGATGAGATCAAGCTGAGAAAGATCAAGCTGAGAGAGGAGGCTGAG 300

875 GAAGATCAAGAGAGAGGATTTGATTTCTGATATCTCTCTTGGCCAAATGAG 934
 301 AAAGATCAAGAGAGAGGATTTGATTTCTGATATCTCTTGAATTTGGCCAAATGAG 360

935 GAATGAGAGATCTTCTGACAGAGAGCTCCGCTGAGTGAAGTGAACCTTATTTGA 994
 361 GAATGAGAGATCTTCTGACAGAGAGCTCCGCTGAGTGAAGTGAACCTTATTTGA 420

995 GGGCCACGACACACAGGAGGAGTCTCTGAGATCTCTATGCTCTGGCCACACACC 1054
 421 GGGCCATGACACACAGGAGGAGTCTCTGAGATCTCTATGCTCTGGCCACACACC 480

1055 CAAGATCAAGAGAGGATGCGGAGAGATCCACAGCTCTCTGAGTGAAGTTCAT 1114
 481 TGATCAACCAAGAGATGAGAGAGAGATTCAGAGCTCTCTGGGAGAGTTCAT 540

1115 CACTTGAACCACTTGAACAGATGCTCTTACACCAACATGATGATTAAGAGGCACTGAG 1174
 541 CACTTGAACCACTTGAACAGATGCTCTTATACCAATGATGATTAAGAGGCACTTGAAG 600

1175 GCTTACCAACCGGTGACAGGATTTGGAGAGGCTCAAGATCCCGTCACTTCCCTGA 1234
 601 GCTTACCAACCGGTGACAGGATTTGGAGAGGCTCAAGATCTCTTCCCTGA 660

1235 TGGGAGCTCTTGGCCAAAGGTATGATGCTCTC-CTTCTCATTTATGAGCTTCAACACA 1293
 661 TGGAGCTCTTTTAACCAAGGTATCCCGGTCAACATCTTCAATTAAGTCTCAACACA 720

1294 ACCCAAAAGTGTGAGCCCAACCAAGAGTGTGTAACCTTTCGTTTGAACCGGGTTC 1351
 721 AACCAAAAGTGTGAGCCCAATTC-GAAGGTTTGAACCTTTCGTTTGAACCGGAATTC 777

RESULT 13
 B1147726 834 bp mRNA linear EST 05-JUN-2001
 LOCUS 602912992F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5053895 5',
 DEFINITION mRNA sequence.
 ACCESSION B1147726
 VERSION B1147726.1 GI:14607727
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 834)
 REFERENCE NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rmail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Place: LAM1146 row: h column: 24
 High quality sequence stop: 831.
 Location/Qualifiers
 1..834
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5053895"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1ib="NCI_CGAP_L19"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
 technology="178 g 220 t"

BASE COUNT 205 a 231 c 178 g 220 t

ORIGIN

Query Match 19.4%; Score 499.2; DB 12; Length 834;
 Best Local Similarity 75.0%; Pred. No. 3.3e-44;
 Matches 624; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

882 AAGAGAGAGGAGGATTTGATTTCTGATATCTCTCTTGGCCAAATGAGATGG 941
 1 AAGAGAGAGGAGGATTTGATTTCTGATATCTCTCTTGGCCAAATGAGATGGA 60

942 AGCATTTGTCAAGACAGGATCTCCGCTGAGTGAAGTGAACCTTCAATGTTTGGGCGAC 1001
 61 AAGAGATTTCTGATTAAGACCTTCGCTGAGTGAAGTGAACCTTCAATGTTTGGGCGAC 120

1002 GACACAGACGAGGAGATCTCTGATCTCTATGCTCTGAGCCACACACCCCAAGCAT 1061
 121 GACACAGACGAGGAGATCTCTGATCTCTATGCTCTGAGAGTGAAGTGAACCTTCAAT 180

1062 CAGAGAGGTGCGGAGAGATCCACAGCTCTCTGAGTGAAGTGAAGCTTCATCCTGG 1121
 181 CAGAGAGGTGAGAGAGATCCAAAGTCTCTGAGAGTGAAGTGAAGCTTCATCCTGG 240

1122 AACCACTTGAACCAAGTCCCTTACACCAACATGATGATTAAGAGGCACTGAGGCTTAC 1181
 241 AATGACCTGAGACAAAGTCCCTTATACCATGATGATTAAGAGGCGCTGAGAGATTCAC 300

1182 CCACCGGTGACAGGATTTGGAGAGGCTCAGACCTCCGCTCACTTCCCTGATGGGCGC 1241
 301 CCTCTGTACCAAGTGTGAGAGGCTCAGACCTCCGCTCACTTCCCTGATGGGCGC 360

1242 TCCCTGCCAAAGGTATGATGCTCTCTCTCAATTAAGGCTTCAACCAAAA 1301
 361 TCCCTGCCAAAGGTATGATGCTCTCTCTCAATTAAGGCTTCAACCAAAA 420

1302 GTGTGAGCCCAACCAAGGATTTGAACCTTTCGTTTGAACCGGCTTCTGCTCAACAC 1361
 421 GTGTGAGCCCAATTCAGAGGATTTGATTCCTTCTGATTTGACACAGGCTTTCCTCCGCGAC 480

1362 AGCCAGCTTCTCTGCTCTTCAAGAGATTAAGAGATGCAATGGGAGAAATTTGGC 1421
 481 AGCCAGCTTCTCTGCTCTTCAAGAGATTAAGAGATGCAATGGGAGAAATTTGGC 540

1422 ATGAAGAGGTGAAGTGGACACGAGGCTGACCTGCTCCGCTTGAAGCTGCTGAT 1481
 541 ATGAAGAGGTGAAGTGGACGCTGAGGCTGACCTGCTCCGCTTGAAGCTGCTGAGAT 600

[illegible]

RESULT 14	LOCUS	DEFINITION	Accession
CB455420	670 bp	mRNA	linear
CB455420	712558	MARCB6 BOV Bos taurus	CDNA 5', mRNA sequence.
			EST 26-MAR-2003

SOURCE ORGANISM	Bos taurus (cow)	Bos taurus
-----------------	------------------	------------

REFERENCE

AUTHORS

COMMENT

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4350
Email: smitchem@marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
plate_PQY8070 row: B column: 10
Seq primer: GTAATACGACTCACTATAGGG.
Location/Qualifiers
1. .670

```

/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH108"
/clone_id="MARCK 6BOV"
/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/embryonic."
BASE COUNT      151 a      199 c      163 g      157 t
ORIGIN

```

Query Match	19.0%	Score 490.4	DB 14	Length 670
Best Local Similarity	83.4%	Pred. No. 3.2e-43		
Matches 557	Conservative	0	Mismatches 111	Indels 0
				Gaps 0

Oy 897 TTGGATTTCTGGATATCTCTCTTGGCCAAATGGAGATCTTGCAGC 956
 Db 2 TTGGACTTCTGGACATCTCTCTCTTGGCAGATGGAGATGGAGCAGCTTGTGAC 61

957 AAGGACCTCCGTCGAGGTGGACAAGTTCATGTTGAAGGGCCAGACACACAGCCAGT 1016
62 GAGGACCTCCGTCGAGGTGGACAAGTTCATGTTGAAGGGTCATGACACCAAGCCAGT 121

[illegible]

Db	122	GGCATCTCCGAAATCTCTAATGCTCTAAGCTTCCCATCTGAGCATACGACAGAGGTGTCCG	181
Qy	1077	GAGAGATTCACAGAGCCTCTTGAGTGAATGAGCCTCATCACTTGAACCACTTGACAG	1136
Db	182	GAAAGATCCAGAGCCTCTGGGGAGATGGGCTCATCACTGGGATCACTGGACAG	241
Qy	1137	ATGCGCTTACACCAATATGTCATTTAAGAGGCACTGAGGCTTACCAACCGGTGCCAGGC	1196
Db	242	ATGCGCTTACACCAATATGTCATTTAAGAGGCAATGAGACTTTATCCACCAATGTCATTC	301
Qy	1197	ATTGGCAGAGATTCAGCACTCCCGTCACTTCCCTGATGGGGCTTCCTTGCCAAAGT	1256
Db	302	ATTGGCAGAGATTCAGCACTCCCGTCACTTCCCTGATGGGGCTTCCTTGCCAAAGT	361
Qy	1257	ATATGAGTCTGCTCTTCATTTATGAGCTTCAACAACAACCAAAAGTGGGCCCAACCA	1316
Db	362	ATCTTATGCTCTCTCTCTTTATGACTTCAACAACCAAAAGTGGGCCCAATCCA	421
Qy	1317	GAGGTGTTTGAACCTTTCCTTTTGAACCGGATTCCTCTCAACACAGCCAGCTTTCCTG	1376
Db	422	GAGGTGTTTGAACCAACCGGATTCCTCAACAGGTTCTACTCAACAGCTATGCTTTCCTG	481
Qy	1377	CCCTTCTCAGAGGATCAAGAACTGCAATTGGGAAAACAATTGCCATGAACGAGCTGAAG	1436
Db	482	CCCTTCTCAGAGGATCAAGAACTGCAATTGGGAAAACAATTGCCATGAATGAGTTGAAA	541
Qy	1437	GTCGACACGGCCCTGAACCTGTCTCCGTTTGAAGCTGTGCTGATCCACAGGATCCCC	1496
Db	542	GTCGACCGTGGCCCTGAACCTGTCTCCGTTTGAAGCTGTGATCCCTCCAGGGTCTCCT	601
Qy	1497	ATGCCCATTCACAGACTGTGTGTTGAATCCAAAATGGAATTCACCTGCGCTCAGAGG	1556
Db	602	GTCGCCCATTCACATCATGTGTGATGCCAAAATGGATTCACCTTGACGCTTCAAGAAA	661
Qy	1557	CTGCCCTAA	1564
Db	662	CTGTCTGA	669

RESULT 15	
Bt143913	
LOCUS	Bt143913
DEFINITION	Bt143913 806 bp mRNA linear EST 05-JUN-2001 602908002.F1 NC1_CGAP_Kid14 Mus musculus cDNA clone IMAGE:506489
ACCESSION	5, mRNA sequence.
VERSION	Bt143913
KEYWORDS	Bt143913 1 GI:14603914
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
	Mus musculus
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (Base 1 to 806) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Straubeberg, Ph.D.

Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
plate: L14M1175 row: C column: 02
High quality sequence stop: 756.

FEATURES	SOURCE	LOCATION/Qualifiers
		1..806
		/organism="Mus musculus"
		/mol_type="mRNA"
		/strain="PvB/N"
		/db_xref="taxon:10090"


```
/clone="IMAGE:5064889"  
/lab_host="DH10B (T1 phage-resistant)"  
/clone_1b="NCI_CGAP_Kid14"  
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI_CGAP library."
BASE COUNT      202 a      215 c      184 g      205 t
ORIGIN
```

```
Query Match      18.7% Score 482.2; DB 12; Length 806;  
Best Local Similarity 77.5%; Pred. No. 2.1e-42;  
Matches 599; Conservative 0; Mismatches 168; Indels 6; Gaps 1;  
  
QY      13 GATCCAGAGAGTCTGACCATGAGTCTCTGTGCTGAGGCCAGAGACTCTGGGTG 72  
DB      1 GATCCAGAGAGTCTGTATGATGATGAGTCTCTGTGCTGAGTCCATCAGATCCAGAGA 60  
QY      73 ATGCTCTGGAATCTCCAGACGGCTCTCCCTGCTATTGCTTCTGTCTGTATCAAG 132  
DB      61 GCATCTCTGAGTACTTCAAGTAGCTCTGTGCTCAGCTGTCTCTCTCTTCAAGA 120  
QY      133 CAGTTGAGCTCTACCTGACAGAGAGTGGCTGTCAAGCCCTGACAGAGTTCCCGTGC 192  
DB      121 CAGCCCACTCTACCTGACAGAGAGTGGCTGTCAAGAGTACTCAGAGATCCCATCC 180  
QY      193 CTCCTCCCACTGGCTCTTGGGCAATCCAGAGCTCAACAGAGCAGAGCTACAC 252  
DB      181 CACTCTCTCACTGGCTCTTGGACACA-----AGATCTTAAGAGCAGAGACTTCAAG 234  
QY      253 GGAATTCAGAAATGGGTGAGAGCATTTCCCAAGTGTCTCTCATTTGCTATGGGAGGCA 312  
DB      235 ATATTCTAATGAGATTAAGAAATTTCCCAAGTGTCTCTCAAGTGTCTGAGGAGAGCA 294  
QY      313 AAGTTGTGTCCAGCTCTATGACCCCTGACCTATTAAGAGTGTCTTGGGAGATCAGACC 372  
DB      295 AAGTGGGCAATTCAGATGATGACCTGACCTACATGAAAGCTGATTTGGGAGATCAGACC 354  
QY      373 CGAATTCAGATGATTTCTTACAGATTTCTGCTCCATGGAATGGGTAAGGCTTGTCTCTGT 432  
DB      355 CAAAGCTTAATGTTCTTACAGATTTCTTACGTTCTCTGATTTGGGCGTGTGCTTCTGC 414  
QY      433 TGAATGGGAGAGCATGGTTCCAGCATGACGAGATGTGACCCAGCTTCCACTATGACA 492  
DB      415 TGGATGAGACAGCATGGTTTCAGACCGAGAGATGTTGACCCAGCTTTCACATATGACA 474  
QY      493 TCCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACAGATGATGCTGACAAATGGG 552  
DB      475 TTCTGAAGCCCTTATACGAAATCATGGCAGACTCTGTGCTGTATGCTGATTAATGGG 534  
QY      553 AAGAGCTCTTGGCAGAGATTTCCCTGAGAGTCTTTACAGCAGTCTCTTGAATGACCC 612  
DB      535 AACAGATTTGGCAGAGATTTCCCTGAGATCTTTTCAGACATCACCTTGAATGACCT 594  
QY      613 TGAACACCATCATGAGTGTGCTTTCAGCCATCAGGGCAGCATCAGGTGACAGAAAT 672  
DB      595 TGAACACCATCATGAGTGTGCTTTCAGCCACAGGGCAGTGTCAAGTTGACAGAAAT 654  
QY      673 CTCAGTCTTACATACAGGCCATTAGTACTGAAACAACCTGTGTTTTTCCCGTGTAGGA 732  
DB      655 ACAAGTCTTATATCAGGCGAGTTGAGGACTGAAAGATCTGTGTTTTTCCCGTGTGCGGA 714  
QY      733 ATGCTTTACCAAGATGACACATCTACAGCTGACCTGCTGGCCGCTGG 785  
DB      715 CATTTTTCACAGAAATGACATCATACAGAGTGTCTTAAATGGCTGACAGG 767
```

Search completed: February 15, 2004, 05:29:04
Job time : 5307 secs